

Cambios conceptuales y metodológicos de los índices de marginación 2020¹

Diana Villasana² y Raúl Romo³

Resumen

El índice de marginación cumple 30 años de estimarse. Este punto sirvió de parteaguas para realizar algunas adecuaciones conceptuales y cambios metodológicos. El más importante y central de todos ellos es que el nuevo método utilizado, el de Distancias Ponderadas al cuadrado, permite la comparación en el tiempo. En segundo término, se encuentra el cálculo por primera vez del índice de marginación urbana por colonia. Esta oportunidad también fue aprovechada para realizar pequeños ajustes en algunos indicadores y adecuaciones a la estratificación por Dalenius & Hodges. Los resultados del censo 2020 representaron información nueva y con todas las posibilidades de desagregación geográfica que se requería.

Términos clave: *Índice de marginación, Análisis de Componentes Principales, Distancias Ponderadas al cuadrado, Dalenius & Hodges, entidad federativa, municipio, localidad, AGEB urbana, colonia urbana.*

Introducción

A partir de la actualización de los índices de marginación, con base en los resultados del censo 2020, se cumplieron 30 años de su estimación. Con la información del censo de 1990, la Secretaría General del Consejo Nacional de Población (SG CONAPO) calculó por primera ocasión el índice de marginación para entidades federativas y municipios, con el

censo de 1995 lo inició para localidades y con el censo de 2000 para áreas geostatísticas básicas (AGEB) urbanas del país.

Incluso antes de la primera publicación del índice de marginación, la SG CONAPO ya había realizado estimaciones que fueron utilizadas para fines internos. La necesidad de contar con indicadores que apoyaran la focalización de acciones y programas de política pública era apremiante.

En este sentido, el método de Análisis de Componentes Principales (ACP) representaba la vanguardia en el avance de la estadística y matemática en el ámbito mundial, por ello fue una elección natural para estimar el índice de marginación. Además, jamás se imaginó el uso generalizado que se haría de él, ni tampoco la demanda que tendría para comparar los resultados en el tiempo.

Tomar la decisión de cambiar la metodología del índice de marginación llevó varios años. En 2012, con el auspicio de la Red de Pobreza y Desarrollo Territorial de CONACYT, se llevó a cabo un seminario donde participaron funcionarios públicos, académicos y miembros de organismos internacionales para intercambiar experiencias. En 2014 se realizó un trabajo que hizo un recuento de varias alternativas: media aritmética, media geométrica, distancias euclidianas, componentes principales mediante la matriz de covarianza, Análisis Factorial Confirmatorio, análisis multivariado mediante conjuntos difusos, modelo de curvas latentes de crecimiento (CONAPO, 2014).

¹ Se agradecen la lectura y comentarios de Rodrigo Jiménez y Eric Barrón.

² Jefa de departamento de Migración Interna, Coordinación de Planeación en Población y Desarrollo, Secretaría General del Consejo Nacional de Población (diana.villasana@conapo.gob.mx).

³ Director de Poblamiento y Desarrollo Regional Sustentable, Coordinación de Planeación en Población y Desarrollo, Secretaría General del Consejo Nacional de Población (raul.romo@conapo.gob.mx).



Dos años después, se publicó el *Índice de marginación por entidad federativa y municipio 2015* (CONAPO, 2016) y, derivado del hecho de que no se pudieron actualizar los cálculos por localidad y AGEB urbana con la Encuesta Intercensal 2015, el equipo técnico responsable dispuso de tiempo para seguir explorando métodos alternativos al ACP.

En 2017 se contrató una consultoría con el Dr. Rubén Hernández Cid, académico del Instituto Tecnológico Autónomo de México, quien dirigió una tesis sobre mejoras al índice de marginación. Como resultado se propuso el método STATIS desarrollado en Francia en 1986 como instrumento de medición longitudinal que permite la comparación en el tiempo (CONAPO y UNFPA, 2017).

Antes de que se levantara el censo 2020 y se publicaran sus primeros resultados a finales del mismo año, el equipo técnico de la SG CONAPO estuvo probando diferentes métodos estadísticos con los resultados de 2015, teniendo en cuenta la mencionada comparabilidad temporal.

Entonces, el 30 aniversario de la estimación del índice de marginación y el censo 2020 en proceso de levantamiento, fueron detonantes principales para emprender cambios conceptuales y metodológicos, que culminaron con la actualización y publicación de las bases de datos y notas técnico-metodológicas para entidades federativas y municipios (CONAPO, 2021a), localidades (CONAPO, 2021b) y AGEB urbanas en 2021 (CONAPO, 2021c), así como de colonias urbanas (2022), realizada por primera ocasión y dada a conocer a inicios del presente año.

En este orden de ideas, el trabajo ejemplifica diversos detalles del índice de marginación por localidad 2020. Se tomó esta decisión por ser una de las desagregaciones geográficas con más variaciones en las categorías por grado,⁴ lo que derivó en una solicitud de información realizada a la institución donde se preguntaban muchas especificidades retomadas aquí.

Cambios conceptuales en el índice de marginación

Los indicadores que se utilizan para la estimación del índice de marginación tuvieron algunos ajustes en el ejercicio 2020. Los cálculos para entidad federativa, municipio y localidad continuaron con el mismo número, nueve para los dos primeros casos y ocho para el tercero; para AGEB urbana aumentaron de diez a once; estos mismos once fueron los que se usaron para colonia urbana (véase esquema 1).

Cinco indicadores sobre vivienda se comparten entre todos los niveles de desagregación geográfica, los tres de servicios (drenaje y excusado, energía eléctrica y agua entubada), el piso de tierra y el de hacinamiento; los restantes son comunes a tres o dos ámbitos geográficos. El uso de indicadores diferentes obedece a su pertinencia, el detalle de escalas pequeñas permite observar una mayor variabilidad en los comportamientos de los datos.

Otro cambio se refiere a la homologación del cálculo en relación con los ocupantes en las viviendas y no en las viviendas de forma directa, para así enfocarse más en la población, lo que aplica para los de servicios y presencia de refrigerador; el indicador de hacinamiento se continuó haciendo respecto a las viviendas en entidad federativa y municipio y con ocupantes para los otros tres ámbitos geográficos. Las recomendaciones nacionales e internacionales sugieren ir en el primer sentido, sin embargo, las magnitudes entre las dos formas de obtención difieren muy poco.

Con los referidos a educación, para la población de 5 años o más sin educación básica se agregó la secundaria, antes solo incluía la primaria, derivado de la obligatoriedad y del incremento de la población que estudia y concluye este nivel de estudios; en los casos del analfabetismo y población de 6 a 14 que no asiste a la escuela, no hubo cambios.

Los indicadores relacionados con población sin derechohabencia a los servicios de salud, población que vive en localidades menores a 5 000 habitantes y población ocupada con ingresos de hasta dos salarios mínimos no presentaron cambios en su cálculo. En tanto, los de ocupantes en viviendas sin internet y ocupantes en viviendas sin celular son nuevos.

⁴ Al tener en cuenta los datos de derivados del censo 2010 (CONAPO, 2012).

Esquema 1. Indicadores utilizados para la estimación del índice de marginación, 2020					
Indicador	Entidad federativa	Municipio	Localidad	AGEB urbana	Colonia urbana
Porcentaje de población de 15 años o más analfabeta (ANALF)	x	x	x		
Porcentaje de población de 15 años o más sin educación básica (SBASC)	x	x	x	x	x
Porcentaje de población de 6 a 14 que no asiste a la escuela				x	x
Porcentaje de población sin derechohabiencia a los servicios de salud				x	x
Porcentaje de ocupantes en viviendas particulares sin drenaje ni excusado (OVSD)	x	x	x	x	x
Porcentaje de ocupantes en viviendas particulares sin energía eléctrica (OVSEE)	x	x	x	x	x
Porcentaje de ocupantes en viviendas particulares sin agua entubada (OVSAE)	x	x	x	x	x
Porcentaje de ocupantes en viviendas particulares con piso de tierra (OVPT)	x	x	x	x	x
Porcentaje de viviendas particulares con hacinamiento (VHAC)	x	x			
Porcentaje de ocupantes en viviendas con hacinamiento (OVHAC)			x	x	x
Porcentaje de población que vive en localidades menores a 5 000 habitantes	x	x			
Porcentaje de población ocupada con ingresos de hasta 2 salarios mínimos	x	x			
Porcentaje de ocupantes en viviendas sin refrigerador (OVSR)			x	x	x
Porcentaje de ocupantes en viviendas sin internet			x	x	x
Porcentaje de ocupantes en viviendas sin celular				x	x
Total	9	9	8	11	11

Fuente: Elaborado por el CONAPO con base en CONAPO (2021a, 2021b, 2021c y 2022).

La aplicación del índice de marginación por localidad

El primer aspecto por resaltar, es el hecho de que hasta la última estimación por localidad, la correspondiente a 2010, en seis de los ocho indicadores utilizados en el índice de marginación se referían a las viviendas particulares habitadas, a diferencia del ejercicio 2020 donde se contemplan a los ocupantes en viviendas particulares habitadas.

Otro cambio fue el ajuste en el cálculo del indicador “Población de 15 años y más sin educación básica”, ya que ahora se contempla la secundaria incompleta o bien su homóloga. Con respecto al indicador “Ocupantes en viviendas particulares habitadas sin drenaje ni excusado”, se consideró que las viviendas no tenían

drenaje si contaban con un desagüe conectado a una tubería que va a dar a una barranca o grieta, o bien, una tubería que va a dar a un río, lago o mar.⁵

Por último, un ajuste de cálculo en el indicador de “Porcentaje de ocupantes en viviendas particulares con algún nivel de hacinamiento”, se construyó como la relación entre la cantidad de personas en la vivienda y alguna aproximación al espacio disponible en la vivienda, considerando que una vivienda presenta hacinamiento si tiene más de 2.5 ocupantes por cuarto dormitorio (INFONAVIT y ONU-HABITAT, 2018).

⁵ El detalle de la metodología puede ser consultado en: https://www.gob.mx/cms/uploads/attachment/file/685308/Nota_tcnica_IML_2020.pdf

Con estos cambios se pensaría que la distribución de los indicadores podría haber cambiado, pero en ambos casos es similar a excepción de la variable “Ocupantes en viviendas particulares con hacinamiento”, que presenta un poco más de variabilidad, pero a pesar de ello, se mantiene la misma tendencia.

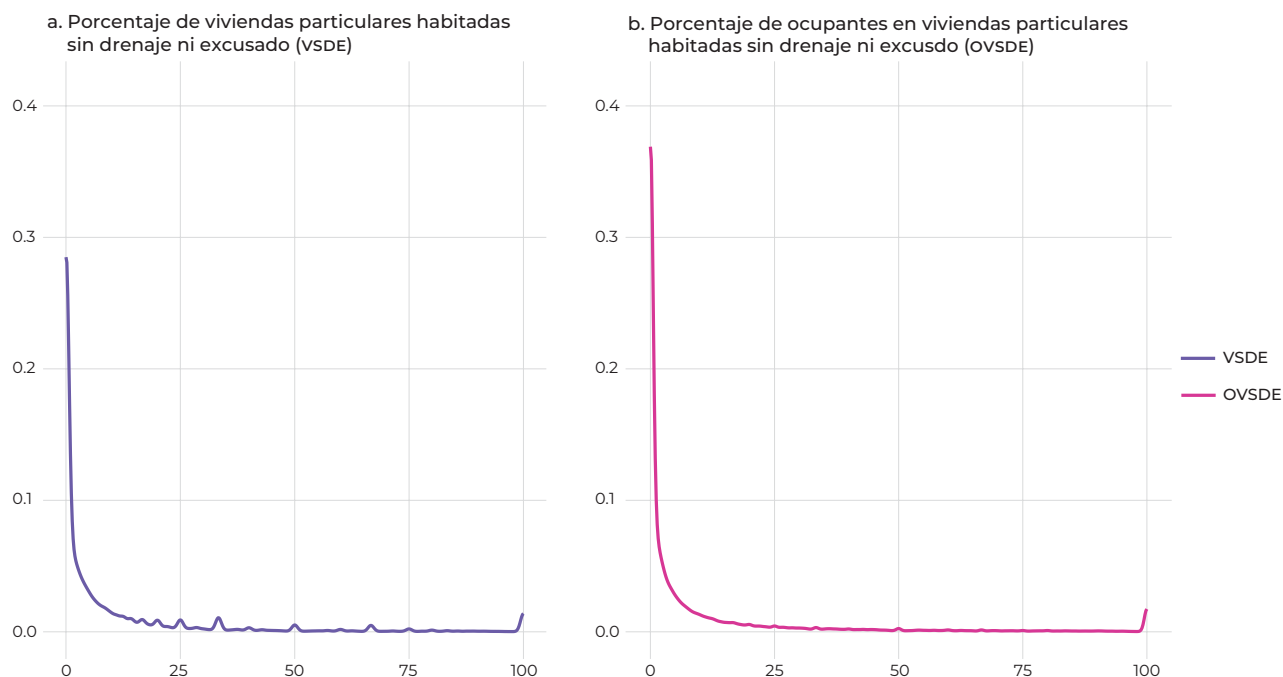
En las gráficas 1-6 se muestran dos escenarios, en el panel izquierdo se presentan los indicadores que contemplan a las viviendas particulares habitadas y en el derecho los ocupantes en viviendas particulares habitadas. La principal diferencia es que las tendencias en las distribuciones de ocupantes en viviendas

particulares son más uniformes (más suavizadas). Por otro lado, estas distribuciones presentan ciertas crestas en determinadas cifras fraccionarias de baja denominación, es decir, que al considerar al total de viviendas en localidades pequeñas recaen en la misma cifra.

Con respecto al cambio de metodología se pueden aclarar algunos aspectos en las secciones siguientes en relación con el ACP y la estatificación de Dalenius y Hodges (1959), las cuales son las principales causas en las discrepancias en el monto de localidades con grados alto y muy alto de marginación entre las estimaciones 2010 y 2020.

Gráfica 1.

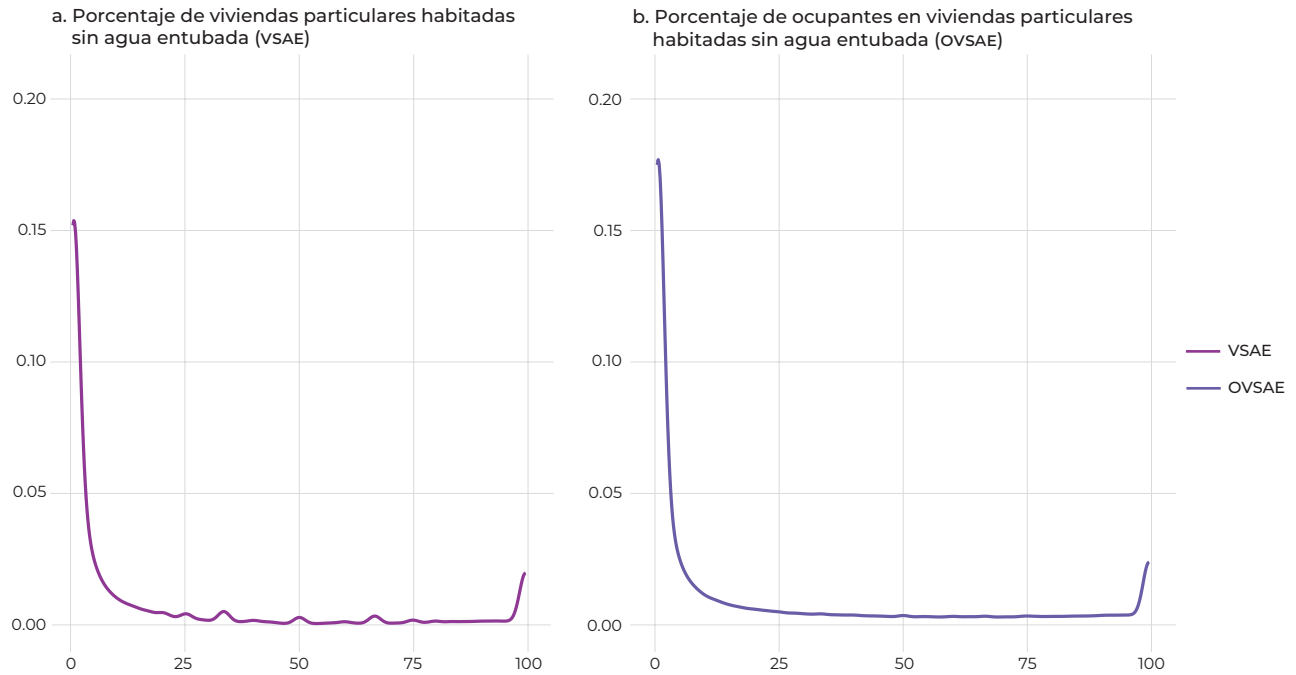
República Mexicana. Distribución de los indicadores sobre drenaje y excusado por localidad, 2020



Fuente: Estimaciones del CONAPO con base en el INEGI, Censo de Población y Vivienda 2020.

Gráfica 2.

República Mexicana. Distribución de los indicadores sobre agua entubada por localidad, 2020



Fuente: Estimaciones del CONAPO con base en el INEGI, Censo de Población y Vivienda 2020.

Gráfica 3.

República Mexicana. Distribución de los indicadores sobre energía eléctrica por localidad, 2020

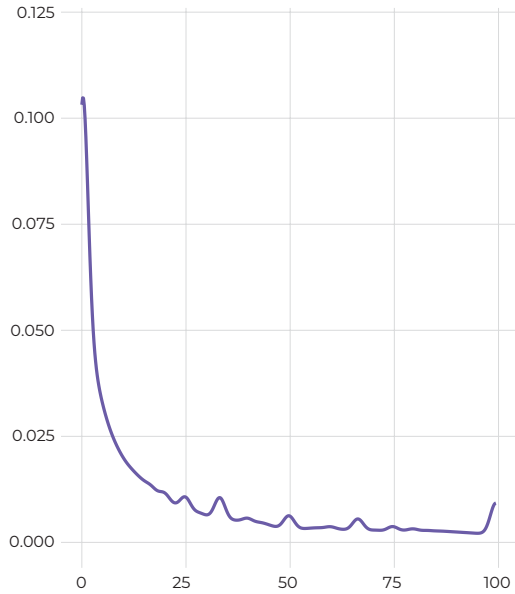


Fuente: Estimaciones del CONAPO con base en el INEGI, Censo de Población y Vivienda 2020.

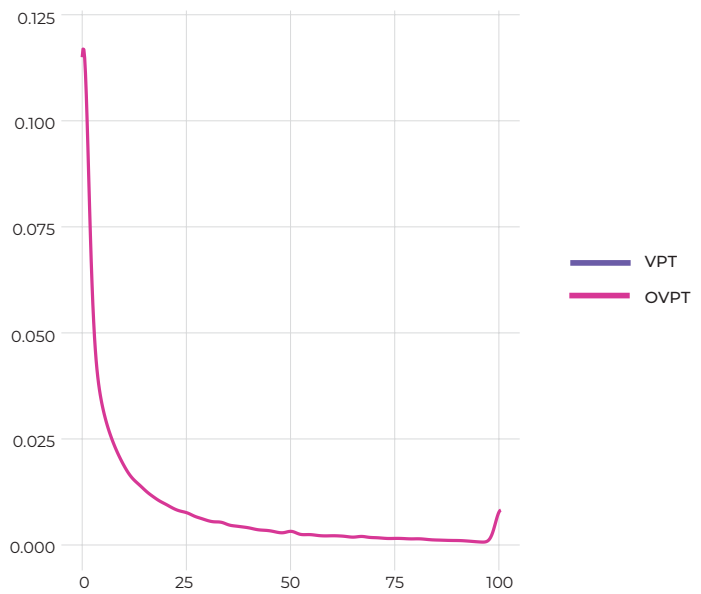
Gráfica 4.

República Mexicana. Distribución de los indicadores sobre piso de tierra por localidad, 2020

a. Porcentaje de viviendas particulares habitadas con piso de tierra (VPT)



b. Porcentaje de ocupantes en viviendas particulares habitadas con piso de tierra (OVPT)

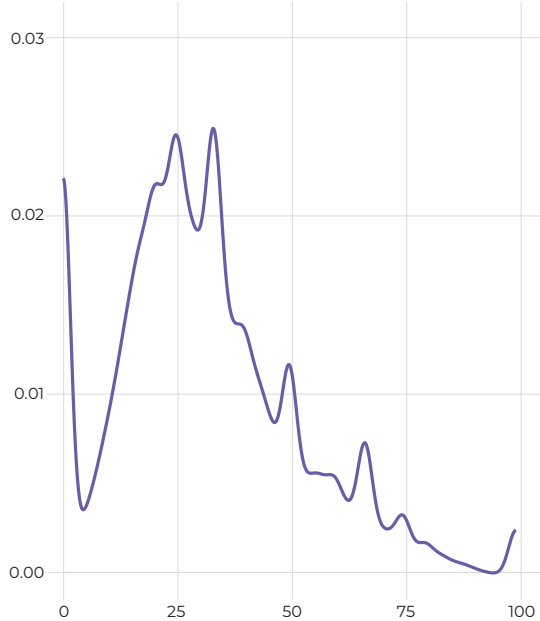


Fuente: Estimaciones del CONAPO con base en el INEGI, Censo de Población y Vivienda 2020.

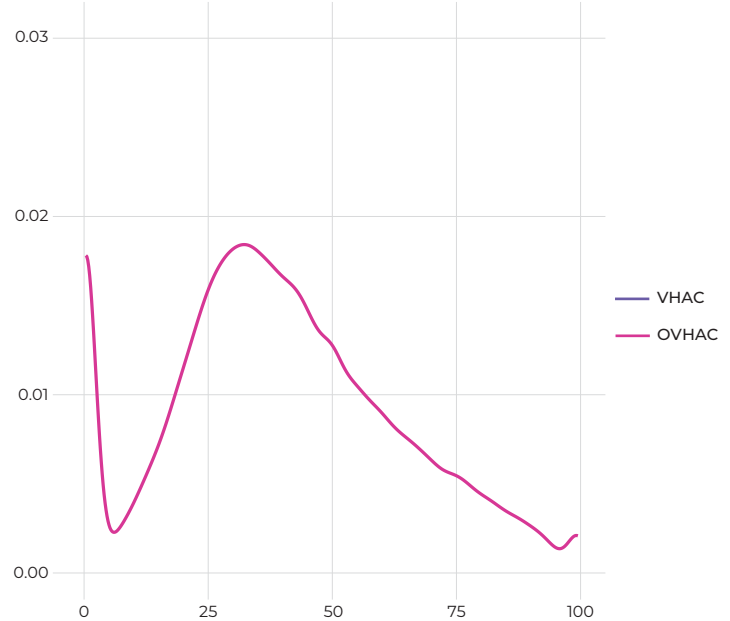
Gráfica 5.

República Mexicana. Distribución de los indicadores sobre hacinamiento por localidad, 2020

a. Porcentaje de viviendas particulares habitadas con algún nivel de hacinamiento (VHAC)



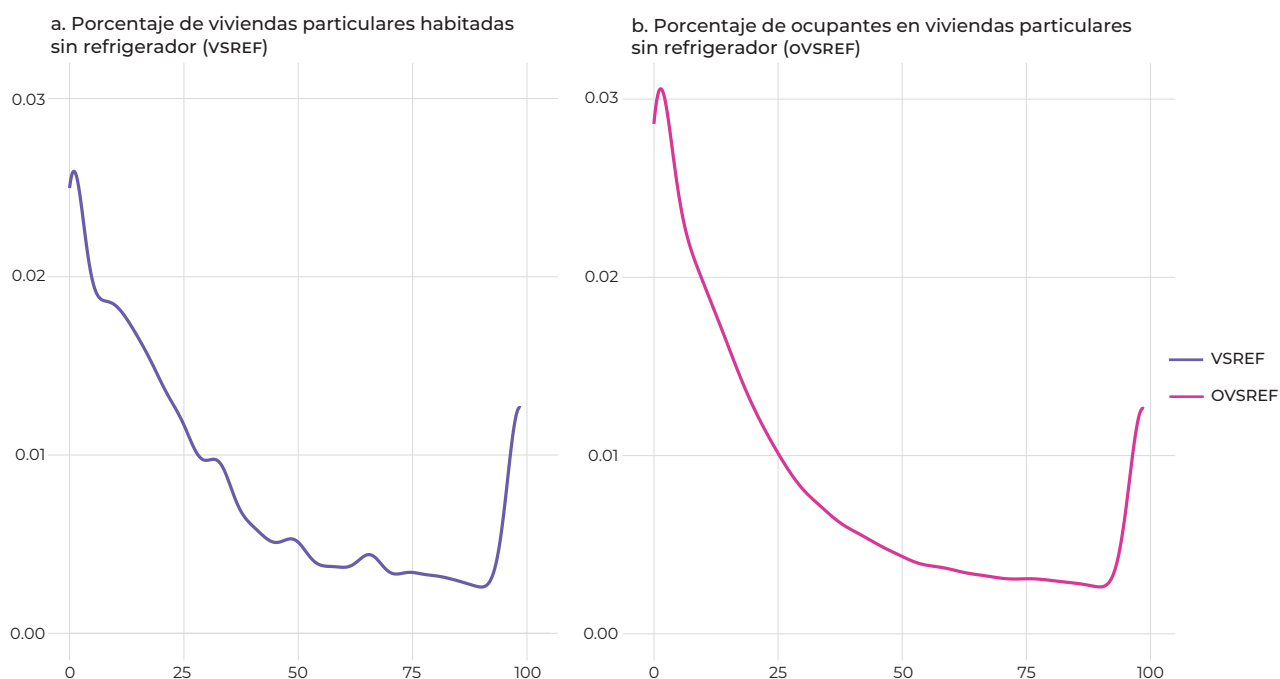
b. Porcentaje de ocupantes en viviendas particulares habitadas con algún nivel de hacinamiento (OVHAC)



Fuente: Estimaciones del CONAPO con base en el INEGI, Censo de Población y Vivienda 2020.

Gráfica 6.

República Mexicana. Distribución de los indicadores sobre presencia de refrigerador por localidad, 2020



Fuente: Estimaciones del CONAPO con base en el INEGI, Censo de Población y Vivienda 2020.

Cambios metodológicos en el índice de marginación

El Análisis de Componentes Principales (ACP)

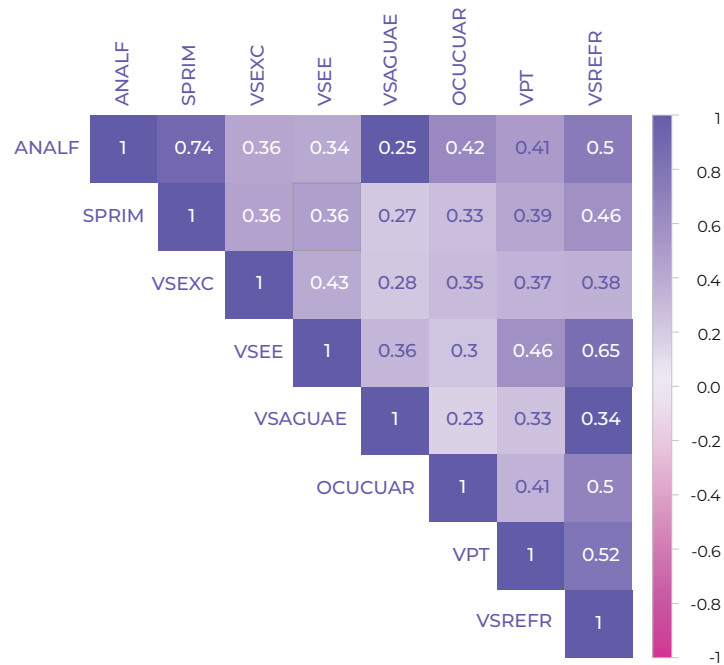
La aplicación de la técnica de ACP tiene como principal fin reducir el número de variables originales sin perder con ello las características del conjunto de datos. Con esto se asegura que la reducción de dimensión sea posible, llevando los indicadores a una medida resumen llamada índice. Esta técnica estadística se basa en la geometría y el álgebra matricial, donde la escritura con matrices logra explicar las bondades del espacio geométrico.

El ACP tiene que cumplir una serie de características para que el modelo sea el adecuado y asegurar una reducción de dimensiones (un índice) y salvaguardar el máximo de información posible. No obstante, con los anteriores índices de marginación ciertas características ya no se cumplían, como se expone a continuación.

La primera componente principal describe la mayor parte de la varianza de los datos, es decir, cuanto más correlacionadas estén las variables originales, mayor es la información explicada en las primeras componentes. Al observar las gráficas 7 y 8 con las matrices de correlaciones, se puede constatar que la interrelación con los indicadores que conforman el índice de marginación por localidad es relativamente baja. Una correlación alta a la hora de interpretar los coeficientes de correlación de Pearson se considera a partir de ± 0.5 a ± 1.0 ; en el caso analizado los indicadores tienen una correlación positiva. Esto indica que la primera componente principal no recogerá el máximo de la información posible y sería necesario contemplar más de una componente para tener un porcentaje adecuado de información, cuestión que no sucede. Es importante recordar que el ACP parte de la matriz de correlaciones, misma que se obtiene de la descomposición matricial de esta.

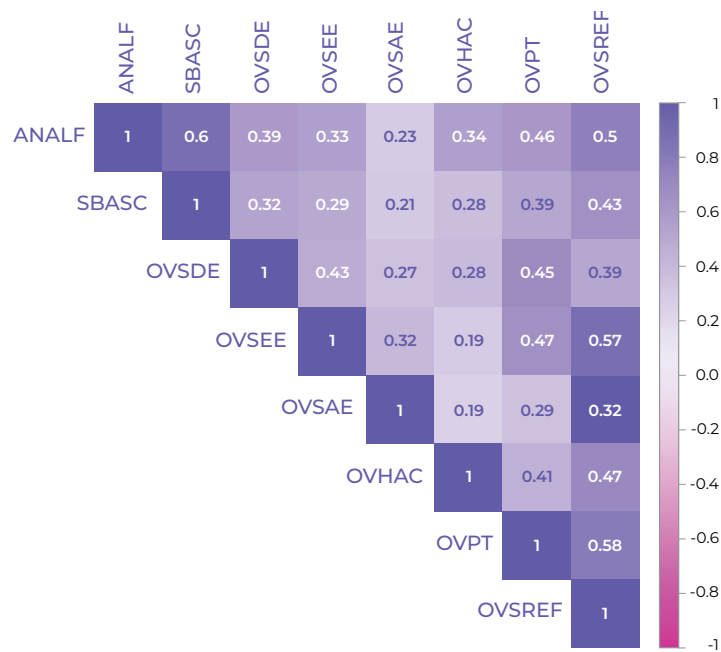
Por otro lado, el método *DP2*, también permite reducir el número de variables en un solo índice (propiedad de unicidad), pero no toma como referencia la

Gráfica 7.
República Mexicana. Matriz de correlaciones, 2010



Fuente: Elaborado por el CONAPO con base en el INEGI, censo 2010.

Gráfica 8.
República Mexicana. Matriz de correlaciones, 2020



Fuente: Elaborado por el CONAPO con base en el INEGI, censo 2020.

matriz de correlación, sino la determinante de la matriz y con ello se evita duplicidad de la información, recogiendo 100 por ciento de la información, algo que con el ACP no era posible.

Otro punto desventajoso del ACP, es que los componentes de bajo orden a veces contienen el aspecto “más importante” de la información y los demás componentes se pueden ignorar. De hecho, en el índice de marginación anterior, solo se tomaba la primera componente principal, perdiéndose así información al no incluir las demás componentes. Al respecto, a continuación se muestran algunos resultados principales que se generaron con dicho método para tratar de responder la pregunta de por qué anteriormente se identificaba una mayor cantidad de localidades con alto y muy alto grado de marginación.

Como se puede apreciar en el cuadro 1, la primera componente principal del índice de marginación por localidad para 2010, recogía 47.8 por ciento de la información, esto es, se tuvo que prescindir de manera parcial de 52 por ciento que aportaban los ocho indicadores. De haberse continuado con la misma metodología, la estimación para 2020 hubiera recogido solo una proporción de 45.8.

La calidad de representación de una variable sobre un plano se evalúa visualmente a partir de la distancia al borde del círculo de radio 1, esto es posible

ya que las variables están centradas y normalizadas a su desviación típica, y a su vez, se proyectan a las componentes tomando como referencia el $\cos^2(\theta)$ y de esta manera poder visualizar la correlación entre una variable y una componente principal (véanse gráficas 9 y 10). Teniendo en cuenta que una representación alta indica una buena correlación de la variable en la componente principal y por otro lado una representación baja indica que la variable no está perfectamente representada. En 2010, el indicador de “Porcentaje de viviendas particulares habitadas sin disponibilidad de agua” se encuentra mejor representado en la tercera componente que en la primera; hay que recordar que el índice de marginación a nivel localidad utilizó la primera componente principal para la realización del índice. Para una variable dada, la suma del $\cos^2(\theta)$ sobre todos los componentes principales será igual a 1.

Este tipo de gráfico se puede hacer para cada una de las observaciones (localidades), pero este tipo de representaciones ocurren en un espacio vectorial diferente al de las variables. No obstante, dada la magnitud de estas, no se apreciarían adecuadamente. Cabe recalcar, como ya se dijo, que esta técnica estadística se basa en la geometría y el álgebra matricial, donde la escritura con matrices logra explicar las bondades del espacio geométrico.

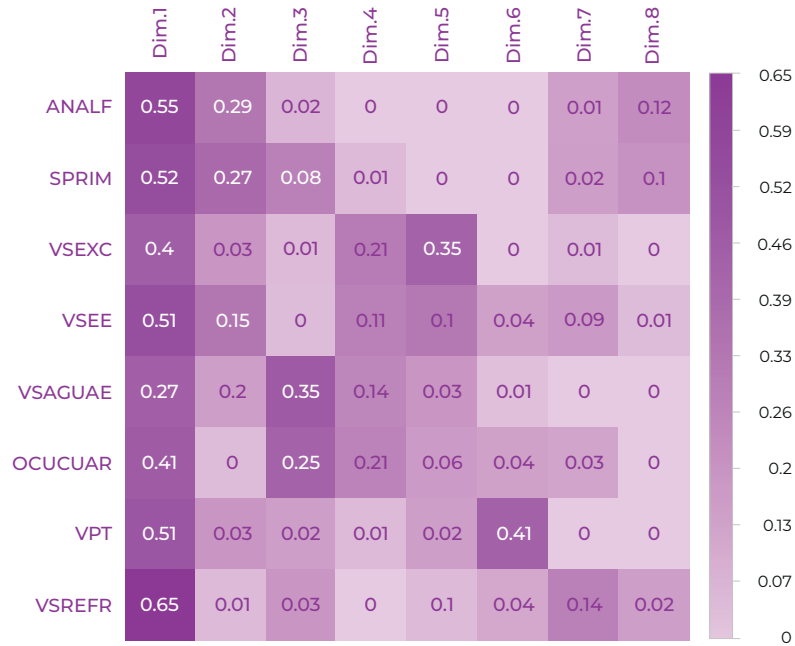
Cuadro 1. República Mexicana. Porcentaje de varianza explicada en el índice de marginación por localidad, 2010 y 2020							
Ejercicio 2010				Ejercicio 2020			
Dimensión	Eigenvalue	Porcentaje de Varianza	Porcentaje de varianza acumulada	Dimensión	Eigenvalue	Porcentaje de Varianza	Porcentaje de varianza acumulada
1	3.82	47.78	47.78	1	3.66	45.80	45.80
2	0.98	12.19	59.97	2	0.95	11.88	57.68
3	0.77	9.59	69.56	3	0.82	10.22	67.90
4	0.68	8.56	78.11	4	0.75	9.42	77.32
5	0.67	8.39	86.50	5	0.64	8.01	85.33
6	0.54	6.75	93.26	6	0.46	5.72	91.05
7	0.29	3.64	96.90	7	0.40	4.96	96.00
8	0.25	3.10	100.00	8	0.32	4.00	100.00

Fuente: Estimaciones del CONAPO con base en el INEGI, censos de población y vivienda 2010 y 2020.



Gráfica 9.

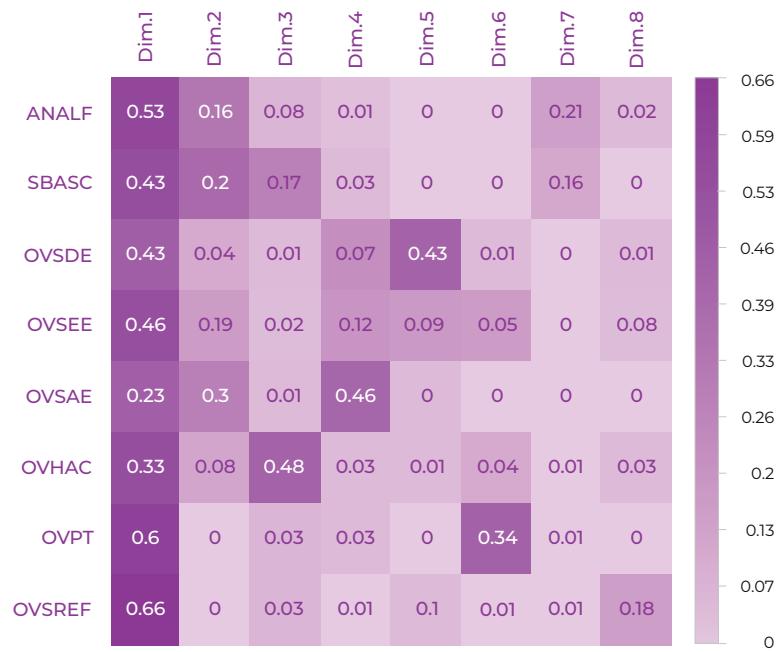
República Mexicana. Contribución de las variables al índice de marginación por localidad, 2010



Fuente: Elaborado por el CONAPO con base en el INEGI, Censo 2010.

Gráfica 10.

República Mexicana. Contribución de las variables al índice de marginación por localidad, 2020



Fuente: Elaborado por el CONAPO con base en el INEGI, censo 2020.

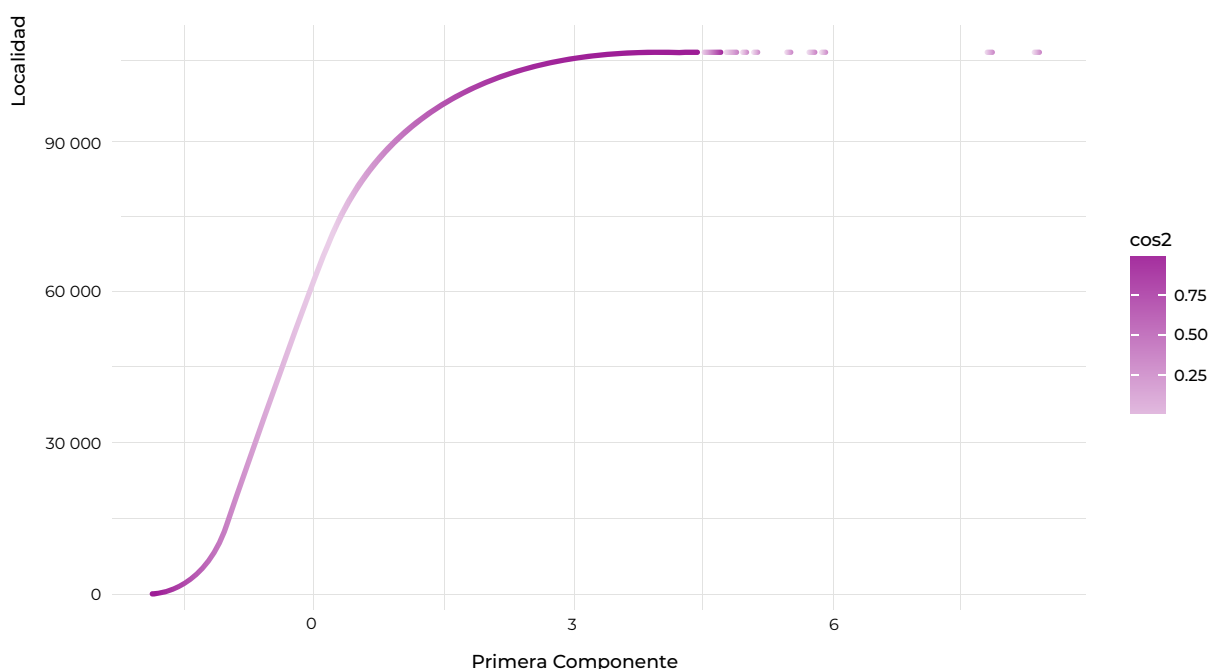
Para apreciar realmente la representación de las observaciones, se mide por el valor cuadrado del coseno $\cos^2(\theta)$ del ángulo del triángulo formado por el punto en el espacio y, su proyección sobre el componente (véanse gráficas 11 y 12). Este tipo de parámetro permite analizar las observaciones que pueden influir y evaluar la estabilidad en la construcción de ciertos componentes, el cual puede tomar valores de 0 a 1, así que a simple vista se puede ver que gran parte de las observaciones tienen contribuciones altas a excepción de la sección central de la gráfica 12. Este tipo de observaciones son explicadas por otra componente principal (la tercera, cuarta, quinta o hasta octava), pero si se ordenan las observaciones de la primera componente (véase gráfica 11), estas se encuentran en la parte central de la distribución. Recordemos que las observaciones en el índice de marginación 2010 presentan valores negativos y positivos, y algunas localidades con marginación muy alta se ubican en la parte positiva y hay localidades con grado muy bajo en la parte negativa.

Con relación a la diferencia en la distribución de las localidades por grado de marginación entre 2010 y 2020, se debe principalmente a la baja representación de algunas localidades en el índice, aunado a que la explicación de la primera componente es muy baja. No existe una respuesta o método único que permita identificar cuál es el número óptimo de componentes principales a utilizar. Con el método *DP2* no existe inconveniente en el porcentaje de información, ya que el factor corrector permite que la información no se repita, cada variable explica información única, y en cierta manera hace que este método sea más robusto con respecto al de ACP.

Tanto la proporción de varianza explicada como la proporción de varianza explicada acumulada, son de gran utilidad a la hora de decidir el número de componentes principales y por lo general el número de componentes seleccionadas se toma cuando el incremento deja de ser sustancial. En la gráfica 13 se puede advertir que después de la segunda componente principal, el comportamiento de las componentes es invariable.

Gráfica 11.

República Mexicana. Representación en la primera componente a nivel localidad, 2010

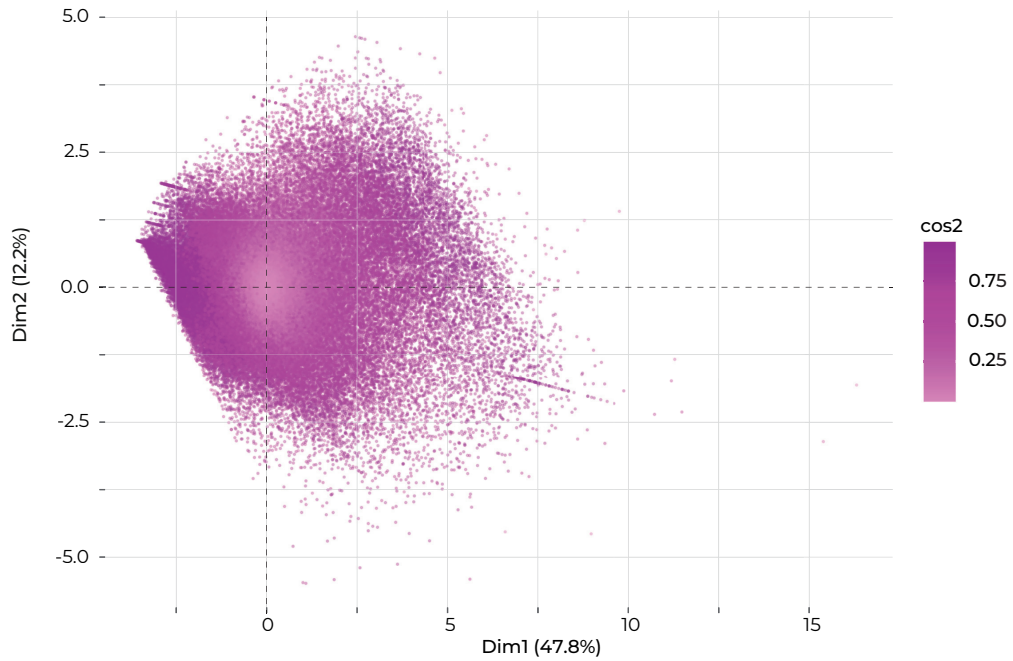


Fuente: Elaborado por el CONAPO con base en el INEGI, censo 2010.



Gráfica 12.

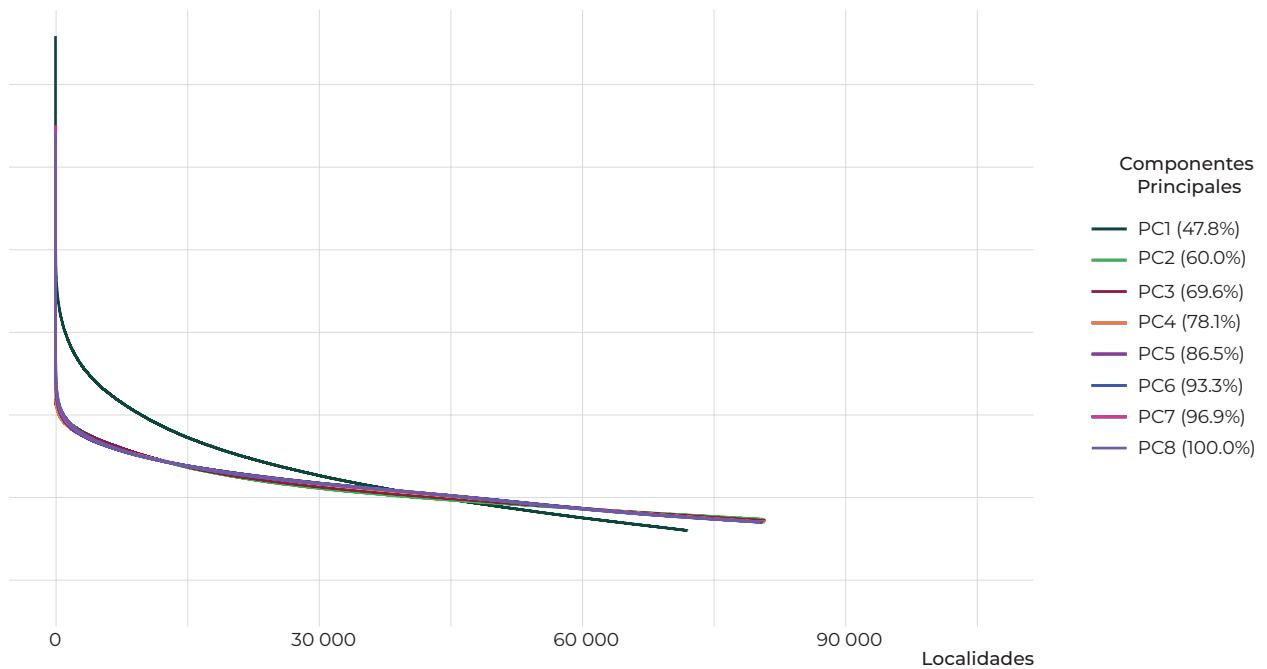
República Mexicana. Representación en las dos primeras componentes a nivel localidad, 2010



Fuente: Elaborado por el CONAPO con base en el INEGI, censo 2010.

Gráfica 13.

República Mexicana. Comportamiento acumulado de las componentes PC1 y PC8 a nivel localidad, 2010



Fuente: Elaborado por el CONAPO con base en el INEGI, censo 2010.

Adicionalmente, y como ya se mencionó, el método *DP2* permite que los resultados sean comparables en el tiempo. Esta cualidad era una demanda de diversas personas interesadas, ante la necesidad de observar los avances/retrocesos conseguidos por los programas y acciones dirigidos a poblaciones y lugares con desventajas sociales. Obviamente que las estimaciones 2020 no serían suficientes para estar en posibilidades de realizar esta tarea, por lo cual se llevó a cabo el cálculo para los datos 2010 y 2015⁶ con la misma metodología. Además, se agregó el índice de marginación normalizado a las bases de datos, que de igual forma, permite vislumbrar la diferencia en las magnitudes del propio índice entre unidades de observación.

Otro asunto a ser resaltado es la estimación por colonia realizada por primera ocasión. Este requerimiento existía por la complejidad de que las personas usuarias ubicaran el AGEB urbana donde residían. De hecho se utilizaron los mismos indicadores que para este último nivel de desagregación geográfica. Aunado a lo anterior y para facilitar aún más las cosas, se está diseñando un sistema de consulta georreferenciado sencillo y práctico con base en R. Por último, también se han puesto a disposición los programas de cálculo o *sintaxis*.

Método de estratificación univariada Dalenius & Hodges

Aunado a lo antes descrito, el método de estratificación también genera discrepancia en la clasificación de los grados de marginación. Anteriormente el índice derivado del ACP solo establecía un orden en las unidades geográficas (medida ordinal) y no era interpretable, se aplicó un método de estratificación univariada para poder agrupar estas observaciones, teniendo siempre como objetivo que la varianza fuera mínima al interior de cada estrato y máxima entre cada uno de ellos, es decir, lo más homogéneos posibles.

En este tipo de método es importante establecer el número de clases en el que se va a segmentar de manera *a priori* la variable de estudio. Considerando que los límites del estrato final dependen de la elección inicial del número de clases y no hay ninguna

teoría que proporcione la mejor, en ejercicios previos se mantenía el mismo criterio sin realizar una evaluación más estricta para determinarlas. Anteriormente, en los índices de marginación por localidad se tomaban como referencia 20 clases, pero dicho criterio no cumplía el objetivo de la estatificación de varianza mínima, y al mismo tiempo, se favorecía a las colas de la distribución.

Por lo anterior, en las estimaciones de 2020 se decidió usar un método iterativo que proporcionara el número de clases y cumplir con el objetivo de que la varianza sea mínima al interior de cada estrato y máxima entre cada uno de ellos.

Para no afectar los límites dados por el método, ya que es muy sensible a la varianza y dado que el número óptimo de clases está basado en la media aritmética, es necesario identificar los datos significativamente distintos al resto de las observaciones, ya que podrían ocasionar errores serios en los cálculos y, por consiguiente, resultados nada fiables. Para contrarrestar este problema se usó el método de caja propuesto por Hubert y Vandervieren (2007), para tomar en cuenta el grado de asimetría de un conjunto de datos, concluyendo que los límites con los que se debe trabajar están dados por:

$$[Q_1 - 1.5 e^{-4MC} IQR; Q_3 + 1.5 e^{3MC} IQR] \text{ para } MC \geq 0$$

$$[Q_1 - 1.5 e^{-3MC} IQR; Q_3 + 1.5 e^{4MC} IQR] \text{ para } MC < 0$$

donde:

Q_1 y Q_3 : hacen referencia al primer y tercer cuartil respectivamente, la diferencia entre estos dos valores da como resultado el valor de espacio intercuartil (*IQR*), y *MC* (*medcouple*) cuantifica el grado de asimetría de una muestra univariable $\{x_1, x_2, \dots, x_n\}$. Este se encuentra definido por la ecuación:

$$MC = \underset{x_i \leq Q_2 \leq x_j}{\text{med}} h(x_i, x_j)$$

donde:

Q_2 : es la mediana de la muestra, y
 x_i y x_j : son elementos de la muestra.

La función *h* está definida por la ecuación:

$$h(x_i, x_j) = \frac{(x_j - Q_2) - (Q_2 - x_i)}{x_j - x_i}, \quad \forall x_i \neq x_j$$

⁶ Considerando la desagregación de los datos, para este año solo se replicaron los resultados por entidad federativa y municipio.



Esto no quiere decir que la identificación de los datos atípicos se haya descartado del modelo, sino que simplemente se identifica la primera observación del primer cuartil y es ahí donde se reposicionan las observaciones que salen de la norma. Con ello se pueden contemplar tales valores en el modelo de estratificación, ya que como su nombre lo dice, es importante medir el número de frecuencias que se encuentra en

cada clase. Con este procedimiento se logró reducir el ruido en la identificación de los límites de los estratos, especialmente, en el “muy alto”.

Una vez que se identificaron los casos atípicos en el índice por localidad y para medir el efecto del número de clases en la varianza de cada estrato, se recurrió al método iterativo para obtener un criterio de agrupación óptimo (véanse cuadros 2 y 3).

Cuadro 2. República Mexicana. Límites para el cálculo de estratificación del índice de marginación por localidad, 2020		
Outliers	Rango	
	Mínimo	Límite
5 364	0	11.7

Fuente: Estimaciones del CONAPO con base en el INEGI, Censo de Población y Vivienda 2020.

Cuadro 3. República Mexicana. Número óptimo de clases del índice de marginación por localidad, 2010 y 2020				
Año	Número de clases	Error estándar	C.V.	Método utilizado
2010	60	0.428	0.029	DP2
2020	23	0.347	0.017	DP2

Fuente: Estimaciones del CONAPO con base en el INEGI, censos de población y vivienda 2010 y 2020.

Consideraciones finales

Cambiar la metodología de estimación del índice de marginación tomó 30 años. En esta decisión se consideraron las demandas de los diversos usuarios y que los ajustes tendrían un impacto, al menos inicial, en la identificación y asignación de recursos a población objetivo de acciones y programas de política pública, lo que requeriría del apoyo técnico necesario para su asimilación; por lo que se sostuvieron varias reuniones con este propósito.

Para algunos estados fue necesario ir más allá, debido a que sus leyes de coordinación fiscal contienen fórmulas para el reparto de presupuesto a sus

municipios, en donde incluyen el valor directo del índice de marginación. Esto fue un problema, al considerar que la metodología anterior generaba índices con números positivos y negativos, y que estos últimos eran las unidades con menor marginación, lo opuesto al resultado 2020.

Para esos casos específicos se estimó una proyección por el método de ortogonalización de Gram-Schmidt, que transforma el índice actual en el anterior. Con ellos se aclaró que es solo una medida provisional en lo que se reforman dichas leyes estatales para cambiar la fórmula para el cálculo de los recursos que se otorgarán a los municipios, ya que el índice publicado en la página web es el oficial.

Existieron múltiples inquietudes de todo tipo que han sido atendidas, reuniones con dependencias federales, estatales y municipales, solicitudes de información de este tipo de instituciones, organismos internacionales, académicos y usuarios en general, presentaciones en comités, grupos de trabajo, seminarios y en sesiones de capacitación interna y a Consejos Estatales de Población y organismos equivalentes. La elaboración de este artículo fue producto de estos trabajos.

En un punto previo a la decisión en el cambio de metodología, se sopesaron las implicaciones que tendría esta medida, incluso las de índole política, no obstante, se privilegió la mejora y transparencia de los procesos técnicos, pues valga decir que por primera ocasión se hizo público el programa de cálculo o sintaxis de programación.

No puede dejar de destacarse la estimación del índice de marginación urbana por colonia. La mayor facilidad con la que un usuario promedio puede identificar la colonia donde reside, y no el AGEB, también fue una demanda realizada desde años previos. Esta decisión fue igualmente valorada y discutida en sentido de la disponibilidad de las capas geográficas.

En términos del método anterior, se recapitula que el ACP retiene poca información en la primera componente principal, lo cual genera cierto sesgo en las colas de la distribución dado que solo un pequeño porcentaje de la información explicaba el fenómeno de la marginación, lo que ocasionó en cierta medida que gran parte de las observaciones cayeran en los grados alto y muy alto.

De la misma forma, es necesario recalcar dos aspectos importantes. En primer término, el índice ya se puede comparar de forma directa. Además, ahora, los números más grandes corresponden a índices de menor marginación. Para estar en la sintonía de la comparabilidad en el tiempo, debieron replicarse los cálculos 2010 y 2015, dichos datos también pueden ser consultados. Por último, se incluye un índice de marginación normalizado, que se representa entre 0 y 1, lo cual permite una comparación numérica más sencilla.

Finalmente, es preciso mencionar que las tres publicaciones que aglutinan las estimaciones para los cinco niveles de desagregación geográfica, con

capítulos de análisis de resultados se encontrarán en la página web de CONAPO.⁷

Bibliografía y fuentes consultadas

- CONAPO [Consejo Nacional de Población]. (2012). *Índice de marginación por localidad 2010*. Recuperado de: https://www.gob.mx/cms/uploads/attachment/file/671850/Indice_de_marginacion_por_localidad_2010.pdf
- _____. (2014). El índice de marginación desde otras perspectivas metodológicas. *La situación demográfica de México 2014*. Recuperado de: <https://www.gob.mx/conapo/documentos/la-situacion-demografica-de-mexico-2014>
- _____. (2016). *Índice de marginación por entidad federativa y municipio 2015*. Recuperado de: <https://www.gob.mx/conapo/documentos/indice-de-marginacion-2015-284579>
- _____. (2021a). Nota técnico-metodológica del Índice de marginación por entidad federativa y municipio 2020. Recuperado de: <https://www.gob.mx/conapo/documentos/indices-de-marginacion-2020-284372>
- _____. (2021b). Nota técnico-metodológica del Índice de marginación por localidad 2020. Recuperado de: <https://www.gob.mx/conapo/documentos/indices-de-marginacion-2020-284372>
- _____. (2021c). Nota técnico-metodológica del Índice de marginación urbana por AGEB 2020. Recuperado de: <https://www.gob.mx/conapo/documentos/indices-de-marginacion-2020-284372>
- _____. (2022). Nota técnico-metodológica del Índice de marginación urbana por colonia 2020. Recuperado de: <https://www.gob.mx/conapo/documentos/indices-de-marginacion-2020-284372>
- _____ y UNFPA [Fondo de Población de las Naciones Unidas]. (2017). *Propuesta conceptual y*

⁷ Estarán disponibles en: <https://www.gob.mx/conapo/documentos/indices-de-marginacion-2020-284372> y se subirán de manera paulatina a partir del segundo semestre de 2022.



metodológica del índice de marginación y análisis de su tendencia en México (mimeo).

Dalenius, T. y Hodges, J. L., Jr. (1959). Minimum variance stratification. *Journal of the American Statistical Association*, 54, 88-101.

Hubert, M. y Vandervieren, E. (2007). An adjusted boxplot for skewed distributions. *ScienceDirect*, 52 (12), 5186-5201. Recuperado de <https://doi.org/10.1016/j.csda.2007.11.008>

Husson, F., Le, S. y Pages, J. (2010). *Exploratory Multivariate Analysis by Example Using R*, Chapman and Hall.

INFONAVIT [Instituto del Fondo Nacional de la Vivienda para los Trabajadores] y ONU-HABITAT [Programa de las Naciones Unidas para los Asentamientos Humanos]. (2018). *Vivienda y ODS en México*. Recuperado de: <https://onuhabitat.org.mx/index.php/la-vivienda-en-elcentro-de-los-ods-en-mexico>