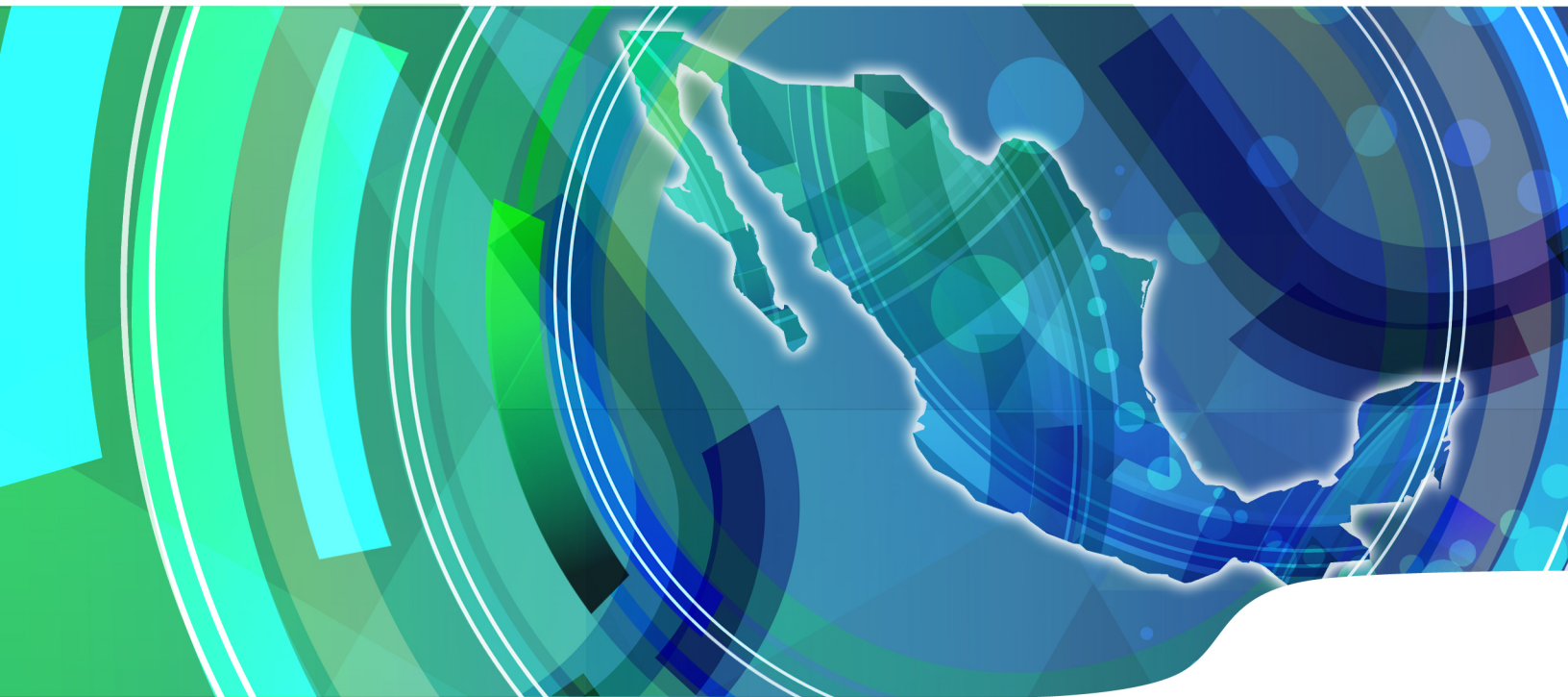


Metodología para la medición de la pobreza en los municipios de México **2010**



Consejo Nacional de Evaluación de la Política de Desarrollo Social**Investigadores académicos****María del Rosario Cárdenas Elizalde**

Universidad Autónoma Metropolitana

Fernando Alberto Cortés Cáceres

El Colegio de México

Agustín Escobar LatapíCentro de Investigaciones y Estudios
Superiores en Antropología Social-Occidente**Salomón Nahmad Sittón**Centro de Investigaciones y Estudios
Superiores en Antropología Social-Pacífico Sur**John Scott Andretta**

Centro de Investigación y Docencia Económicas

Graciela María Teruel Belismelis

Universidad Iberoamericana

Secretaría Ejecutiva**Gonzalo Hernández Licona**

Secretario Ejecutivo

Ricardo César Aparicio Jiménez

Director General Adjunto de Análisis de la Pobreza

Thania Paola de la Garza Navarrete

Directora General Adjunta de Evaluación

Edgar A. Martínez Mendoza

Director General Adjunto de Coordinación

Daniel Gutiérrez Cruz

Director General Adjunto de Administración

Colaboradores

Equipo técnico

Sergio Ulises Andraca Castillo
María del Rosario Cárdenas Elizalde
Fernando Alberto Cortés Cáceres

David Ricardo Escamilla Guerrero
Gerardo Antonio Escaroz Cetina
Jorge Fernando González Moreno
Carlos Mora Jurado
Martha Moreno Pérez
Diana Negrete González
Nayeli Noyolitzin Salgado Granados
Norma Alejandra Vergara Lope Gracia
Ricardo Aparicio

El CONEVAL reconoce la valiosa colaboración del equipo técnico para el proyecto de estimaciones en áreas pequeñas coordinado por El Colegio de México, conformado por los investigadores nacionales, internacionales y asistentes de investigación, Delfino Vargas, Jorge de la Vega, Nicolaos Tzavidis, Luis Enrique Nieto, Gonzalo Pérez, Christian Carmona y Jesús Luján. Se reconocen también las propuestas metodológicas y la participación en los seminarios de áreas pequeñas de Ignacio Méndez, Hortensia Moreno, Christopher Brunsdon, Jae Kwang-Kim y Rafael Pérez-Abreu. Asimismo, reconocemos las valiosas sugerencias aportadas por Rosa María Rubalcava, Enrique Minor, Cristina Pérez, Rodrigo Aranda y Martín Lima. También agradecemos al equipo técnico que inició el proyecto de las estimaciones municipales en el año 2009, Víctor Pérez, Dulce Cano y Rocío Espinosa.

Contenido

Introducción	7
1. Planteamiento del problema y exploración de metodologías	8
1.1 Planteamiento del problema.....	8
1.2 Exploración de metodologías	9
1.3 Implementación de metodologías	10
2. Selección de la metodología	11
2.1 Criterios generales de selección	12
2.2 Criterios de precisión de las estimaciones.....	13
2.3 Contraste de métodos	14
3. Construcción de estratos.....	16
4. Indicadores de carencia social.....	20
4.1 Indicadores de carencia social estimados directamente	20
4.2 Construcción del indicador de carencia de acceso a la seguridad social.....	20
4.3 Construcción del indicador de carencia de acceso a la alimentación.....	22
5. Indicadores relacionados con el espacio del bienestar económico.....	23
6. Integración de los indicadores de pobreza (ELLG) y calibración de la población	25
6.1 Integración de los indicadores de pobreza (metodología ELLG).....	25
6.2 Calibración de las estimaciones de pobreza.....	26
7. Profundidad e intensidad de la pobreza.....	27
8. Indicadores de cohesión social	28
Conclusiones	29
A. Anexo matemático.....	35
A.1 Resumen.....	35
A.1.1 Identificación y notación básica	36
A.1.2 Agregación	38
A.2 Método ELL para la estimación del ingreso.....	40
A.3 Método ELLG para la estimación de las medidas de incidencia, profundidad e intensidad de la pobreza, así como los indicadores de cohesión social	44
A.4 Modelos de regresión logística usados para la estimación de los indicadores de carencias por acceso a la seguridad social y por acceso a la alimentación	51

A.4.1 Modelo de regresión logística para el indicador de carencia por acceso a la alimentación	51
A.4.2 Modelo de regresión logística mixta para el indicador de carencia por acceso a la seguridad social.....	52
A.5 Construcción del Índice de Privación Social (<i>IPS</i>).....	53
A.6 Calibración de los factores de expansión de la muestra del censo.....	54
B. Mapa conceptual del proceso de investigación	58
B.1 Etapa de exploración de metodologías	58
B.2 Implementación de metodologías y cálculo de la medición de pobreza nacional y por entidad federativa.....	59
B.3 Contraste de resultados, selección de la metodología y presentación de resultados	60
C. Información complementaria	61

Introducción

La Ley General de Desarrollo Social (LGDS) mandata al Consejo Nacional de Evaluación de la Política de Desarrollo Social (CONEVAL) a realizar la medición de la pobreza de forma multidimensional y considerar, al menos, los indicadores de ingreso corriente per cápita, rezago educativo, acceso a los servicios de salud, acceso a la seguridad social, calidad y espacios de la vivienda, acceso a los servicios básicos en la vivienda, acceso a la alimentación y el grado de cohesión social. De acuerdo con la ley, este ejercicio debe ser realizado cada dos años a escala nacional y estatal, y cada cinco a escala municipal. Para ello se debe utilizar información proveniente del Instituto Nacional de Estadística y Geografía (INEGI).

Con el objetivo de dar cumplimiento a sus obligaciones legales, el CONEVAL, en estrecha colaboración con el INEGI, desarrolló el Módulo de Condiciones Socioeconómicas de la Encuesta Nacional de Ingresos y Gastos de los Hogares (MCS-ENIGH). Este módulo ha permitido realizar la medición de la pobreza en México a escala nacional y estatal para los años 2008 y 2010. El MCS-ENIGH ha sido un instrumento fundamental no solamente para conocer el porcentaje de la población en situación de pobreza, sino también para observar los niveles de carencia en cada uno de los indicadores mencionados en la Metodología para la Medición Multidimensional de la Pobreza en México (CONEVAL, 2010).¹

De acuerdo con esta metodología, son tres los espacios analíticos relevantes para el estudio de la pobreza: el del bienestar económico, el de los derechos sociales y un espacio territorial. Mediante los dos primeros se calcula la incidencia, profundidad e intensidad de la pobreza en México. El tercer espacio permite clasificar a las entidades de acuerdo con su grado de cohesión social.

El objetivo de este documento metodológico es describir los procesos de investigación, desarrollo, evaluación, implementación y cálculo que se llevaron a cabo para realizar la medición de la pobreza a escala municipal para México correspondientes al año 2010.²

El proceso de investigación que siguió el CONEVAL se puede esquematizar en las siguientes cinco etapas: planteamiento del problema; exploración de metodologías; implementación de metodologías; contraste de resultados y selección de la metodología.

En la primera sección del documento se describe el problema de la medición de la pobreza en los municipios del país y se describen, de forma general, las distintas metodologías que se exploraron para solucionarlo. En la segunda sección se aborda la selección de la metodología. Se describen los criterios generales y específicos que debían cumplir las distintas opciones metodológicas, se revisan los resultados de cada una de ellas y se analiza si los resultados ofrecían soluciones acordes a los criterios establecidos previamente por el Consejo. Las secciones subsecuentes dan cuenta de las fases que se siguieron al implementar la metodología seleccionada para obtener las estimaciones de pobreza en los municipios del país. En la tercera sección se describe el proceso

¹ Los resultados obtenidos por el CONEVAL para la medición de la pobreza y sus dimensiones a escala nacional y por entidad federativa en los años 2008 y 2010 están disponibles en la página electrónica www.coneval.gob.mx.

² Para un lector no familiarizado con la medición de la pobreza en México se le recomienda consultar la Metodología para la Medición Multidimensional de la Pobreza en México (CONEVAL, 2010). El presente documento se basa en los indicadores de pobreza definidos en la misma.

de construcción de estratos de municipios. En la cuarta sección se describe cómo se hicieron las estimaciones relacionadas con las carencias sociales. Los indicadores relacionados con el ingreso se describen en la quinta sección. El proceso para integrar los componentes de la pobreza y obtener los indicadores relacionados con los cuadrantes de pobreza (pobreza, pobreza extrema, pobreza moderada, población vulnerable y población no pobre y no vulnerable) se desarrolla en la sexta sección; como un apartado de la misma sección, también se describe cómo, a través de una técnica de calibración estadística, se consiguió la consistencia entre las estimaciones municipales y por entidad federativa. En la séptima sección se describe la construcción de la medida de profundidad en el espacio de los derechos sociales. Los indicadores relacionados con la cohesión social se describen en la octava sección. Finalmente, se ofrecen algunas reflexiones y conclusiones generales y, en tres anexos, se presentan el desarrollo matemático de la metodología implementada, un mapa conceptual que ilustra de manera general el proceso metodológico, así como información complementaria relevante.

1. Planteamiento del problema y exploración de metodologías

En esta sección se abordarán el planteamiento del problema que implicó el desarrollo de la medición municipal y la exploración de las metodologías realizada en 2009 y 2010.

1.1 Planteamiento del problema

El cálculo de la pobreza a escala municipal en México representó un desafío desde diferentes puntos de vista. En primer lugar, el CONEVAL enfrentó una serie de restricciones en cuanto a las fuentes de información. De acuerdo con la LGDS, las fuentes que deben utilizarse son aquellas publicadas por el INEGI.

Las fuentes de información con las que contó el CONEVAL para medir la pobreza a escala municipal fueron dos. Por un lado, el MCS-ENIGH, que contiene toda la información necesaria para medir la pobreza a niveles estatal y nacional. Sin embargo, con la información del MCS-ENIGH no es posible realizar estimaciones precisas a escala municipal de la pobreza y sus indicadores tal como lo marcan la LGDS y los lineamientos y criterios generales para la definición, identificación y medición de la pobreza.³

Por otra parte, también se disponía del Censo de Población y Vivienda 2010. El censo proporciona información válida a escala municipal de la población nacional pero con ella no es posible calcular todos los indicadores de carencia social y de ingreso. Asimismo, asociada al censo se levantó una muestra con un cuestionario ampliado. La muestra posibilita la estimación directa y precisa de cuatro de los indicadores de pobreza a escala municipal: el rezago educativo, el acceso a los servicios de salud, la calidad y espacios de la vivienda y el acceso a los servicios básicos en la vivienda.⁴

³ Los lineamientos y criterios generales para la definición, identificación y medición de la pobreza fueron publicados por el CONEVAL en el Diario Oficial de la Federación el 16 de junio de 2010.

⁴ La muestra del Censo de Población y Vivienda 2010 consta de 2,903,604 viviendas. Su diseño permite considerar como dominio de estudio a cada uno de los 2,456 municipios que había en el país en 2010.

Si bien fueron calculados de forma directa estos cuatro indicadores, la muestra del censo no permite generar toda la información necesaria para la estimación de la pobreza como la ha definido el CONEVAL. La información que proporciona la muestra del censo es insuficiente para calcular los indicadores de acceso a la seguridad social, de acceso a la alimentación y el ingreso corriente total per cápita (ICTPC).

En México no existe una fuente de información que permita estimar la pobreza a escala municipal de manera directa, tal como se calcula la pobreza nacional y estatal con el MCS-ENIGH. Ante tal situación, el CONEVAL, en su calidad de institución encargada de definir los lineamientos y criterios para la definición, identificación y medición de la pobreza, se dio a la tarea de explorar y analizar un conjunto de metodologías existentes en el contexto internacional mediante las cuales fuera posible estimar los indicadores faltantes para calcular la pobreza a escala municipal.

En la aplicación de estas metodologías se buscó no sólo dar cumplimiento a los requerimientos establecidos por la LGDS sino también que los resultados municipales fueran consistentes con los obtenidos a escala nacional y estatal que se habían reportado previamente (el 29 de julio de 2011).⁵ Adicionalmente, el método seleccionado, de acuerdo con los principios que rigen el trabajo del CONEVAL, debía garantizar la transparencia, la objetividad y el rigor técnico de la metodología adoptada.

Para obtener los mejores resultados, el CONEVAL llevó a cabo un proceso de investigación que se dividió en tres etapas: exploración de metodologías, durante los años de 2009 y 2010; implementación de metodologías, en los años 2010 y 2011, y finalmente, la selección de la metodología de estimación definitiva y su instrumentación. La presentación de los resultados se llevó a cabo el dos de diciembre de 2011.⁶

1.2 Exploración de metodologías

Durante los años 2009 y 2010 se consultó a un conjunto de expertos nacionales e internacionales y se promovieron seminarios y talleres de trabajo sobre temas estadísticos de estimación en áreas pequeñas. Los modelos de estimación en áreas pequeñas tienen como objetivo resolver el problema de disponer de estimadores confiables, válidos y estadísticamente precisos cuando no se cuenta con tamaños de muestra suficientemente grandes en todas las áreas geográficas en las cuales se divide a la población. En el caso de la medición municipal, las áreas pequeñas son los municipios del país. Para generar una solución al problema planteado se construyen modelos estadísticos que combinan fuentes de información distintas y se estiman de manera indirecta los indicadores de interés.

En particular, se firmó un convenio con El Colegio de México, cuyo objetivo fue que distintos investigadores nacionales e internacionales propusieran metodologías para estimar de manera

⁵ La fuente de información para los indicadores nacionales y por entidad federativa es el MCS-ENIGH 2010. Para la estimación de pobreza municipal, se combinan la muestra del Censo y el MCS-ENIGH, por lo cual las estimaciones no necesariamente son consistentes entre los municipios y las entidades federativas. Para resolver este problema se implementó un método de calibración que se tratará a detalle en la sección 6.2.

⁶ Los resultados de la medición de la pobreza en los municipios del país se pueden consultar en la página del CONEVAL www.coneval.gob.mx

indirecta los indicadores de pobreza en los municipios del país.⁷ Asimismo, el CONEVAL requirió evaluar los resultados de las metodologías con la información disponible, para contar con una selección previa de modelos a implementar de acuerdo a la evaluación de sus resultados. Los modelos se implementaron con la información censal del año 2000, así como de la Encuesta Nacional de Ingresos y Gastos de los Hogares (ENIGH) de ese mismo año.⁸ En esa etapa de exploración se realizaron pruebas con los datos de una entidad federativa. El estado que se seleccionó fue Veracruz por ser el único con sobre-muestra⁹ para ese año (1,580 hogares, consultar el documento metodológico de la ENIGH 2000), además de contar con población rural y urbana, municipios con alta y con baja incidencia de pobreza, con población indígena y no indígena; en otras palabras, se trata de una entidad que refleja la diversidad económica y social de la población mexicana.

Los modelos propuestos por los distintos académicos fueron los siguientes: el modelo híbrido, propuesto por el Dr. Jae Kwan Kim; el de estimadores Horwitz-Thompson, el cual estuvo precedido por un análisis de conglomerados, implementado por el Dr. Ignacio Méndez Ramírez y la Dra. Hortensia Moreno Macías; el modelo bayesiano (MB), desarrollado por el Dr. Luis Enrique Nieto; los Modelos Cuantil-M y del mejor predictor empírico (MPE), propuestos por el Dr. Nikos Tzavidiz, y el modelo de regresión geográficamente ponderado, propuesto por el Dr. Chris Brunsdon.¹⁰ No es objeto de este documento explicar a detalle los modelos explorados; no obstante, los lectores interesados en las propuestas pueden consultar las referencias pertinentes en la bibliografía.

1.3 Implementación de metodologías

Al analizar los resultados de las metodologías propuestas por los investigadores (consultar el cuadro C.1 del anexo C), se decidió implementar, en una segunda etapa, —con todos los municipios del país y con información de 2010—, las metodologías del MPE y el MB.¹¹ El MPE permitió obtener una estimación robusta de los parámetros basados en la variable (continua) ingreso; además, proporcionó errores cuadráticos medios pequeños; otra de sus ventajas es que podía combinarse con el enfoque del modelo de regresión geográficamente ponderado para incluir también información socioespacial. Por su parte, el MB permitió una estimación multivariada para variables continuas y dicotómicas; también pudo utilizarse la información geográfica y fue muy flexible: se puede adaptar fácilmente a modelos en las treinta y dos entidades federativas o algún otro grupo de municipios. Además de las metodologías implementadas, se incorporó la idea de formar grupos de municipios homogéneos, propuesta por los doctores Ignacio Méndez y Hortensia Moreno, con la finalidad de encontrar variables determinantes o explicativas de la pobreza de acuerdo con el perfil social y económico de los distintos grupos de municipios y para obtener mayor

⁷ Consultar la Metodología para la medición de la pobreza en México (CONEVAL 2010) y el anexo A de este documento.

⁸ La información más reciente era la del Censo de Población y Vivienda y la ENIGH, ambos del año 2005; sin embargo, el censo se realizó con un cuestionario con un menor número de preguntas respecto al cuestionario del censo para el año 2000. Toda vez que se requería que la estructura de las fuentes de información fuera similar a la del censo del año 2010, se determinó utilizar la información del año 2000.

⁹ El término sobre-muestra hace referencia a que el tamaño de muestra en la encuesta es suficiente para considerar como dominio de estudio a la entidad federativa.

¹⁰ Consultar las metodologías en: Metodología de ajuste e imputación de indicadores de la pobreza por ingresos en áreas pequeñas (Vargas y colaboradores, 2010).

¹¹ Este proceso se llevó a cabo durante 2010 y 2011.

precisión en las estimaciones, suponiendo que entre mas homogéneos fuesen los grupos se obtendría una menor varianza y, por lo tanto, mayor precisión de los estimadores.

La implementación y prueba de las metodologías se realizó de octubre de 2010 a noviembre de 2011. El CONEVAL, además de las dos metodologías propuestas por los investigadores, también estimó de manera interna modelos de regresión logística jerárquicos (modelos de regresión logística mixtos) y una combinación de un modelo multinomial para las categorías de ingreso y de modelos de regresión logística jerárquicos para el acceso a la alimentación y el acceso a la seguridad social.¹² Como se tratará a detalle en la segunda sección, ninguno de estos modelos satisfizo todos los criterios generales de selección, razón por la cual el CONEVAL se vio en la necesidad de implementar el método del mejor predictor empírico generalizado a la medición multidimensional de la pobreza (MPEG)¹³ y el método de Elbers y colaboradores generalizado a la medición multidimensional de la pobreza (ELLG, por las siglas de sus autores Elbers, Lanjouw y Lanjouw). Estas metodologías se implementaron de manera independiente, y si bien los modelos comparten muchas de las variables determinantes, éstas no necesariamente son las mismas en ambos métodos debido a las particularidades de cada uno de los modelos y a que los implementaron diferentes grupos de trabajo.

2. Selección de la metodología

En la presente sección se comentan los criterios definidos por el CONEVAL para seleccionar la metodología de medición de la pobreza en los municipios del país; asimismo se presenta un contraste de las metodologías que se implementaron. Cabe mencionar que la selección de las metodologías se realizó a finales de 2011, cuando se tuvieron los resultados de todos los métodos implementados, por lo que esta sección servirá de síntesis del trabajo que se realizó para la medición de la pobreza en los municipios del país. En las secciones subsecuentes se detallará el proceso que siguió el Consejo para implementar la metodología seleccionada, sin abundar en los detalles técnicos de las metodologías que no fueron incorporadas por las razones que se expondrán en esta sección.

Para llevar a cabo la selección de la metodología fue necesario considerar varios aspectos. Por un lado, existen algunas propiedades de la medición de la pobreza que son necesarias para cumplir con los ordenamientos de la LGDS y con la metodología de medición oficial de pobreza. Les hemos denominado criterios generales de selección. La metodología debía permitir reportar también los indicadores de incidencia (proporción de la población), profundidad e intensidad de la pobreza, según se especifican en el apartado 2.1. Asimismo, se requería que la información fuera consistente al momento de desagregarla tanto por dimensiones de la pobreza como por grupos de población.

¹² Los modelos de regresión logística jerárquicos o mixtos se pueden consultar en las secciones del anexo matemático A.4.1 y A.4.2. (Hoffman, 2004).

¹³ En la aplicación del método MPEG, la estimación del ICTPC se realizó como una aplicación del método MPE (Vargas y colaboradores 2010) y se combinó con las estimaciones directas y sintéticas de los indicadores de carencias sociales tal como se realizó con el método de Elbers y colaboradores generalizado a la medición de la pobreza, el cual que se desarrolla en la sección seis.

Otros criterios deseables, que llamaremos de precisión, se refieren a qué tan próximos, en términos cuantitativos, eran las estimaciones que se generaron con los métodos implementados de las estimaciones por entidad federativa que se obtuvieron directamente del MCS-ENIGH 2010 y que el CONEVAL ya había reportado previamente.

2.1 Criterios generales de selección

La metodología de estimación municipal de la pobreza debía cumplir con las siguientes cuatro propiedades:

Incidencia, profundidad e intensidad de la pobreza a escala municipal

Debía permitir calcular la proporción de la población que es pobre, así como la profundidad y la intensidad de la pobreza (CONEVAL, 2010). La incidencia se mide a través del cociente que resulta de relacionar el número de personas en esta condición y el número total de personas de una población específica; en este caso, la población de interés se refiere a la residente en los distintos municipios del país. La profundidad se define como la proporción promedio de carencias sociales de las poblaciones en pobreza y pobreza extrema. La intensidad de la pobreza se conceptúa como la proporción de carencias de la población pobre respecto del máximo posible de privaciones de la población total y se puede calcular como el producto de la incidencia y la profundidad de la pobreza. De este modo, si un método permitía calcular las dos primeras medidas entonces también permitiría conocer la intensidad de la pobreza en los municipios.

Desagregación de los indicadores por grupos de población

De acuerdo con la metodología oficial, los pasos para la medición de la pobreza consisten en lo siguiente: primero, identificar a la población en situación de carencia y, después, agregar esta población al nivel o grupo poblacional requerido. Por ejemplo, si se cuenta con cualquier indicador por municipio, el indicador por entidad federativa consistirá en la agregación de los municipios que la conformen. Un proceso similar aplica para estimar la pobreza de cualquier subgrupo de la población (*v. gr.* hombres, mujeres; indígenas, no indígenas etcétera).

Desagregación de los indicadores por dimensiones de la pobreza

De manera análoga al criterio anterior, la agregación de las personas identificadas con algún indicador en particular no sólo permite medir la pobreza sino también la contribución de cada una de las dimensiones que la componen. De esta forma, la metodología propuesta debía permitir que la suma de las incidencias de las diferentes carencias fuese consistente con los resultados agregados de pobreza. Por ejemplo, considérese el indicador de calidad y espacios de la vivienda, que, a su vez, está compuesto por cuatro indicadores: la carencia por material de pisos, la carencia por material de muros, la carencia por material de techos y el índice de hacinamiento. Las estimaciones son consistentes si al agregar estos cuatro indicadores como se señala en la metodología para la medición multidimensional de la pobreza, se obtiene el indicador agregado de calidad y espacios de la vivienda.

Medición integrada de la cohesión social

El tercer espacio de análisis incluido en la medición de pobreza, además del espacio de los derechos sociales y del espacio del bienestar, es el relativo al contexto territorial. De acuerdo con la metodología de medición de la pobreza, esta dimensión se determina a partir de los indicadores de cohesión social (CONEVAL, 2010). Entre las medidas que conforman este espacio se encuentran el índice de Gini y la razón del ingreso de la población pobre extrema entre el ingreso de la población no pobre y no vulnerable. Ambos indicadores pueden ser obtenidos a partir del conocimiento de la distribución del ingreso de la población, lo cual permite conocer también la heterogeneidad y concentración del ingreso en los municipios.

Cabe señalar que es posible obtener la estimación de los indicadores de cohesión social de manera independiente; sin embargo, es esencial que su estimación esté integrada a la medición de la pobreza, con la finalidad de que los resultados obtenidos y, en específico, los correspondientes al espacio del bienestar económico (ingreso), sean consistentes con los resultados calculados para la cohesión social (CONEVAL, 2010)

2.2 Criterios de precisión de las estimaciones

Además de las propiedades generales, era necesario considerar también algunos criterios adicionales sobre la consistencia de las estimaciones municipales respecto a las de las entidades federativas. Como se ha comentado, se contaba con dos fuentes de información del año 2010 (la muestra del censo y el MCS-ENIGH). Se deseaba que las estimaciones de los modelos tuvieran una precisión al menos similar a la que se obtenía al contrastar las dos fuentes de información, para lo cual se compararon las estimaciones por entidad federativa de los indicadores de rezago educativo, acceso a los servicios de salud, calidad y espacios de la vivienda, y servicios básicos en la vivienda (los indicadores que se pueden calcular directamente en ambas muestras). El cuadro 2.1 contiene la información resumen de las estadísticas que se utilizaron para definir los criterios (el detalle por entidad federativa se puede consultar en los cuadros C.2 y C.3 del anexo C).¹⁴

Cuadro 2.1
Indicadores de la precisión entre la muestra del Censo de Población y Vivienda 2010 y el MCS-ENIGH 2010, según carencia social

Indicador	Rezago educativo	Acceso a los servicios de salud	Calidad y espacios de la vivienda	Servicios básicos en la vivienda	Promedio
Número de entidades federativas de la muestra del censo que están dentro de los intervalos de confianza del MCS-ENIGH	22	20	20	18	20
Promedio de las diferencias absolutas entre los estimadores puntuales de la muestra del censo y el MCS-ENIGH de las entidades federativas	1.4	2.8	2.2	3.5	2.5
Diferencia absoluta máxima entre los estimadores puntuales de la muestra del censo y el MCS-ENIGH de las entidades federativas	3.8	8.6	5.9	10.8	7.3

Fuente: Estimaciones del CONEVAL con base en la muestra del Censo de Población y Vivienda 2010 y el MCS-ENIGH 2010

¹⁴ Como se mencionó en la sección 1.1, las estimaciones de pobreza que se calcularon con el MCS-ENIGH 2010 se habían publicado previamente. Además, la serie por entidades federativas se genera de manera bianual, razón por la cual fue necesario calibrar las estimaciones municipales a dichos resultados.

Fueron dos los criterios que el CONEVAL definió para las estimaciones indirectas: el primero consistió en comparar el número de entidades federativas para las cuales las estimaciones de la muestra del censo estuvieran comprendidas en los intervalos de confianza de las estimaciones de pobreza estatal calculadas con el MCS-ENIGH 2010. Se estipuló que al menos veintiún estados de la República estuvieran dentro de los intervalos de confianza, mejorando la precisión promedio entre las fuentes de información (20 entidades federativas en promedio). El segundo criterio corresponde a la comparación del tamaño promedio de las diferencias entre las estimaciones mencionadas (estatales del MCS-ENIGH 2010 y estimaciones indirectas), las cuales no debían ser mayores a tres puntos porcentuales. Nótese que este criterio es menor, en todos los casos, que la diferencia máxima observada.

Ambos criterios debían cumplirse para las estimaciones que se calcularon con las diferentes metodologías propuestas. Se contrastaron los resultados para los siguientes seis indicadores: pobreza, pobreza extrema, ingreso inferior a la línea de bienestar (LB), ingreso inferior a la línea de bienestar mínimo (LBM), acceso a la seguridad social y acceso a la alimentación, que eran los que debían ser estimados de forma indirecta.

2.3 Contraste de métodos

El cuadro 2.2 presenta de manera esquemática la comparación realizada de los resultados entre las distintas metodologías de estimación de la pobreza municipal implementadas por el CONEVAL, de acuerdo con los criterios generales y de precisión mencionados anteriormente.

Cuadro 2.2
Comparativo de los resultados de las metodologías de medición multidimensional de la pobreza en los municipios del país, según los criterios de selección. México, 2010

Criterios de selección	Modelo del mejor predictor empírico	Modelo bayesiano	Modelos logísticos jerárquicos ¹	Modelo multinomial para el ingreso y logísticos jerárquicos para carencias sociales ²	Generalización del método del mejor predictor empírico a la medición multidimensional de la pobreza	Generalización del método de Elbers y colaboradores a la medición multidimensional de la pobreza
Criterios metodológicos generales						
Posibilidad de calcular:						
La incidencia de la pobreza (porcentaje de la población)	Sí	Sí	Sí	Sí	Sí	Sí
La profundidad de la pobreza (Proporción promedio de carencias sociales)	No	Sí	Sí	Sí	Sí	Sí
Desagregación por población	No	No	Sí	Sí	Sí	Sí
Desagregación por dimensiones de pobreza	No	Sí	Sí	Sí	Sí	Sí
Medición integrada de la cohesión social	No	No	No	No	Sí	Sí
Precisión de las estimaciones						
Número de entidades para las cuales la estimación de los modelos está comprendida en el intervalo de confianza del MCS-ENIGH 2010 (Debe ser mayor o igual que 21)						
Pobreza	26	10	23	24	18	26
Pobreza extrema	31	20	26	23	1	26
Ingreso menor que la LB	30	12	22	25	11	23
Ingreso menor que la LBM	26	17	26	26	2	22
Carencia por acceso a la seguridad social	28	11	25	28	25	26
Carencia por acceso a la alimentación	22	16	20	27	20	24
	6 de 6	0 de 6	6 de 6	6 de 6	1 de 6	6 de 6
Tamaño promedio de las diferencias entre la estimación de los modelos y el MCS-ENIGH 2010 (Debe ser menor o igual que 3 puntos porcentuales)						
Pobreza	2.2	8.1	2.4	2.6	4.0	2.6
Pobreza extrema	0.8	3.0	1.1	1.3	16.6	1.2
Ingreso menor que la LB	2.5	7.8	3.1	3.1	7.9	3.2
Ingreso menor que la LBM	1.7	3.8	1.6	1.9	13.3	2.4
Carencia por acceso a la seguridad social	1.6	4.5	2.1	2.0	2.1	2.0
Carencia por acceso a la alimentación	2.9	3.7	3.3	2.3	3.3	2.6
	6 de 6	1 de 6	4 de 6	5 de 6	1 de 6	5 de 6

Nota: Los modelos se implementaron con la información de la muestra del Censo de Población y vivienda 2010 y el MCS-ENIGH 2010.

¹ Se realizaron modelos logísticos jerárquicos para los indicadores relacionados con el ingreso, el acceso a la seguridad social y el acceso a la alimentación.

² Para los dos estimadores asociados al ingreso se implementó un modelo multinomial que dividía a la población en tres grupos (con ingreso inferior a la LBM, entre la LBM e inferior a la LB e ingreso mayor o igual a la LB). Los modelos logísticos jerárquicos para el acceso a la seguridad social y el acceso a la alimentación se estimaron con modelos de regresión logística.

Como se puede apreciar en el cuadro, el método MPE fue el más preciso. No obstante, no permitía calcular los indicadores correspondientes a la cohesión social, ni los indicadores asociados a la profundidad de la pobreza y no resolvía el problema de identificación de los individuos, con su repercusión en la desagregación por poblaciones. Respecto a la cohesión social, el Dr. Tzavidis propuso incorporar medidas de entropía, las cuales técnicamente generaban medidas alternativas respecto a la distribución del ingreso y, por lo tanto, eran una alternativa robusta para estimar el grado de la cohesión social, pero sus resultados no correspondían a los indicadores establecidos en la metodología oficial de medición de la pobreza, de acuerdo con los lineamientos y criterios generales para la definición, identificación y medición de la pobreza (CONEVAL, 2010). En cuanto a la profundidad de la pobreza, para la estimación de proporción promedio de las carencias sociales se implementó un proceso *Poisson*¹⁵ que estimaba las carencias sociales y se contaba con un indicador del promedio de carencias en los municipios, lo cual, si bien solventaba el

¹⁵ Para mayor detalle se sugiere consultar Kingman (1993).

indicador a escala municipal, no arrojaba resultados consistentes con las carencias estimadas a nivel individual.

El MB tenía la ventaja de permitir calcular los indicadores como una función de distribución conjunta de un vector multidimensional, por lo que se estiman simultáneamente todos los indicadores. Sin embargo, dado que la distribución se estimó a nivel municipal, no era desagregable para distintos subgrupos de interés. El Dr. Nieto propuso definir las poblaciones de interés y construir una función de distribución conjunta para cada una de éstas; no obstante, este procedimiento generaba inconsistencias al agregar las poblaciones a un nivel superior. Por otra parte, en términos de los resultados, las estimaciones del modelo Bayesiano tuvieron menor convergencia respecto a la metodología del MPE, por lo que el Dr. Nieto sugirió tomar como estimador puntual la mediana, con lo que las estimaciones mejoraban marginalmente.

Los modelos logísticos jerárquicos y el modelo multinomial tampoco cumplían con todos los criterios generales. Por esta razón, el CONEVAL debió implementar el MPEG y el ELLG, ambos métodos cumplían estos criterios. Éstos métodos se basan en los modelos del MPE y el ELL respectivamente para estimar el ingreso y a partir de éste se integran los componentes faltantes de la pobreza (lo cual se desarrolla a detalle en las secciones subsecuentes). Para implementar el MPEG y el ELLG se desarrollaron dos etapas que se describen a continuación.

En una primera etapa se compararon los resultados de los errores absolutos relativos (ARE por sus siglas en inglés, Rao, 2003) de las estimaciones del ingreso de las entidades federativas con respecto a las estimaciones del MCS-ENIGH 2010 realizadas con el método del MPE por el equipo de El Colegio de México y el ELL por el equipo del CONEVAL.

En el cuadro C.4 del anexo C se muestra que los resultados estatales de las estimaciones de la población con ingreso inferior a la línea de bienestar en ambos métodos presentaron errores absolutos relativos similares; sin embargo al comparar los resultados estatales de las estimaciones de población con ingreso inferior a la línea de bienestar mínimo, los errores absolutos relativos en la aplicación del método MPE fueron mayores que los de la aplicación del método ELL.

A partir de los resultados anteriores se calculó el promedio de los ARE. La última parte del cuadro C.4 muestra que en el ELL (10.3) el error promedio fue menor que en el MPE (20.0), lo cual indica que, en este ejercicio, el ELL resultó ser un método más preciso.

En una segunda etapa se generalizaron los modelos MPE (MPEG) y ELL (ELLG) a fin incorporar el componente de carencias sociales en la estimación multidimensional de la pobreza. Con estos dos modelos se contrastó la precisión de la medición multidimensional de la pobreza.¹⁶

Con el ELLG se obtuvieron mejores resultados respecto al MPEG (como se mencionó en la sección 1.3 las variables determinantes de ambos modelos no son las mismas) en todos los criterios de precisión. Los resultados para el ELLG estuvieron dentro de los criterios fijados para la precisión de los estimadores (cumplió con once de los doce criterios), salvo en el caso del

¹⁶ La metodología del MPE constituye, entre otros elementos, el resultado de una revisión crítica al método ELL. Si bien en esta aplicación el MPEG no produjo los resultados esperados, sus fortalezas teóricas sugieren la pertinencia de continuar profundizando la línea de investigación asociada con este método.

indicador de población con ingreso inferior a la línea de bienestar, para el cual las diferencias, en promedio, fueron de 3.2 puntos porcentuales cuando se pretendía que fueran menores o iguales a tres puntos porcentuales. Aun así, la diferencia fue menor que la del MPEG (de 7.9 puntos).

Dado que la implementación empírica del ELLG resultó ser mejor de acuerdo con los criterios fijados previamente por el CONEVAL, se seleccionó como la metodología para estimar la pobreza en los municipios de México para el año 2010. En las secciones subsecuentes y en el anexo matemático se detallará la metodología implementada.

3. Construcción de estratos

Una de las ideas surgidas en el proceso de medición de la pobreza, particularmente durante la exploración de las metodologías, fue la construcción de estratos de municipios. La premisa era que, dada la heterogeneidad de los municipios en el país, así como la diversidad de factores que determinan la pobreza, resulta pertinente agrupar municipios con particularidades comunes. La idea es agrupar municipios en los cuales los determinantes de la pobreza son similares y que, a la vez, difieren de manera significativa de otros grupos de municipios en los factores condicionantes de la pobreza.

Para la construcción de estratos se probaron dos métodos: el de análisis de conglomerados (Mardia y colaboradores, 1997) y el de análisis de clases latentes (Muthén y Muthén, 2010). El objetivo de probar dos metodologías era contrastar los resultados y evaluar las ventajas y desventajas de cada una. Sin pretender dar una explicación exhaustiva, en los dos siguientes párrafos se abordará un breve análisis de ambas técnicas.

El análisis de conglomerados es una técnica que se basa principalmente en criterios matemáticos. Divide a los grupos de acuerdo a la distancia vectorial (mediante una norma prefijada) entre cada observación multivariada y el centro de cada grupo. Cabe mencionar que el número de grupos se determina previamente y luego se construyen los centros de grupo y cada observación. En nuestro caso, cada municipio queda en el grupo en el cual la distancia a su centro sea menor.

El análisis de clases latentes (ACL) es una técnica estadística, la cual se utiliza cuando se busca medir algún fenómeno que no es directamente observable. Para este caso particular, se buscaba representar la variable latente pobreza. El ACL se utilizó para representar la heterogeneidad no observable en un conjunto de municipios (Hagenaars y McCutcheon, 2002; Muthén y Muthén, 2010). A partir de una serie de características agregadas a nivel municipal, el ACL asigna a cada municipio una probabilidad de pertenecer a algún grupo (clase) de municipios similares entre sí. La cantidad de clases no está definida *a priori* sino que se determina mediante un modelo estadístico que especifica la distribución conjunta de las variables para establecer la cantidad de grupos más adecuada, lo cual sucede cuando se minimizan las diferencias al interior de cada grupo y se maximizan las diferencias entre ellos.

Teóricamente, se prefería la metodología del ACL por las ventajas de no tener que definir *a priori* el número de clases y porque se obtenía para cada observación la probabilidad de pertenencia a cada grupo. Sin embargo, al evaluar los resultados de ambas técnicas, se observó que, en general, la clasificación era bastante similar: más de 85 por ciento de los municipios fueron clasificados en

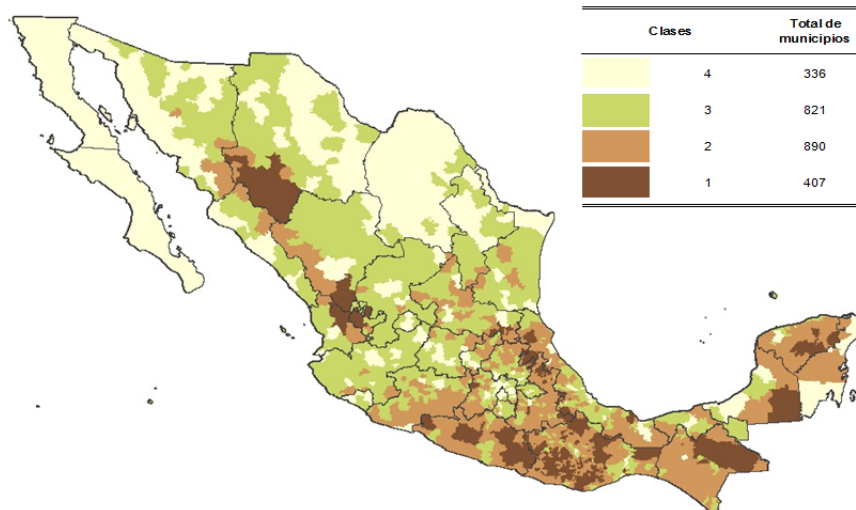
los mismos grupos con ambas técnicas. Dadas sus ventajas, se decidió implementar la clasificación de los municipios en estratos con la técnica de ACL, teniendo como respaldo que ambas técnicas tenían un buen nivel de convergencia.

Las variables utilizadas para el análisis fueron diecinueve (cuadro C.5 del anexo C); todas estas variables pertenecen a las diversas dimensiones que identifican si una persona es pobre según la LGDS: ingreso, educación, salud, seguridad social, vivienda y alimentación.

Las variables fueron agregadas a escala municipal y se utilizó la información más reciente disponible al momento de realizar la estratificación. Con excepción de las relativas al ingreso y la alimentación, la información corresponde a 2010. Además, se incorporó una variable de presencia de población hablante de lengua indígena en los municipios, pues los resultados de mediciones previas han puesto de manifiesto la alta incidencia de pobreza para la población con esta condición. Para las variables relativas al ingreso y a la alimentación se utilizó la información más reciente disponible al momento de la construcción de los estratos de municipios, la cual correspondía a datos de 2005 para el ingreso y de 2006 para los indicadores nutricionales.

El número de clases latentes se definió a partir de encontrar los criterios técnicos siguientes: menor índice BIC-Ajustado; menor valor de la prueba de Lo-Mendel Rubin; mayor nivel de entropía; porcentaje de clasificación mayor al cinco por cierto, y probabilidad de pertenencia a cada grupo mayor de 0.9.¹⁷ Con cuatro clases latentes (estratos) se satisficieron de la mejor manera estos criterios (cuadro C.6 del anexo C). En el mapa 3.1 se presenta la distribución de los municipios de acuerdo con la estratificación realizada.

Mapa 3.1
Distribución de los municipios del país por estrato socioeconómico
(análisis de clases latentes). México, 2010.



Fuente: estimaciones del CONEVAL con base en la muestra del Censo de Población y Vivienda 2010, el Censo de Población y Vivienda 2005, la ENIGH 2005, y los Mapas de Nutrición en México 2005-2006.

¹⁷ Los detalles matemáticos de los criterios citados para la construcción de estratos se pueden consultar en Muthén y Muthén (2010).

El estrato uno contiene a los 407 municipios con mayores carencias y está conformado por 4.6 por ciento de la población nacional. En el otro extremo, se encuentra el estrato cuatro, formado por los 336 municipios en mejores condiciones socioeconómicas, mayoritariamente urbanos, y donde reside cerca de 60 por ciento de la población mexicana (cuadro 3.1).

Cuadro 3.1**Distribución de los municipios y su población
de acuerdo con el análisis de clases latentes**

Estrato	Número de personas	Porcentaje	Número de municipios
1	5,137,875	4.6	407
2	14,569,100	13.0	890
3	26,198,412	23.4	823
4	65,938,397	59.0	336
Total	111,843,784	100.0	2,456

Fuente: Elaboración del CONEVAL con información de la Muestra del Censo de Población y Vivienda 2010.

La estratificación permite tener una clara diferenciación de los niveles de pobreza y pobreza extrema de los cuatro grupos de municipios (cuadro 3.2). En el estrato uno, la incidencia de la pobreza fue de 88.2 por ciento; en el dos, de 74.2 por ciento; en el tres, de 57.8 por ciento, y en el cuatro de 33.7 por ciento. En cuanto a la pobreza extrema, los porcentajes de población en dicha situación para cada estrato fueron, respectivamente, de 54.4, 28.9, 12.7 y 4.4 por ciento. Esta clasificación de municipios no solamente permite una estimación de la pobreza más precisa al interior de cada estrato, sino que puede constituir un instrumento importante para el diseño y aplicación de políticas públicas diferenciadas, las cuales potencien la posibilidad de obtener resultados más eficaces al tomar en cuenta la necesidad de atender poblaciones heterogéneas.

Cuadro 3.2
Porcentaje de la población en situación de pobreza,
según estrato socioeconómico, México, 2010.

Indicadores de pobreza	Estrato			
	1	2	3	4
Pobreza				
Población en situación de pobreza	88.2	74.2	57.8	33.7
Población en situación de pobreza extrema	54.4	28.9	12.7	4.4
Población no pobre multidimensional y no vulnerable	1.4	3.9	10.2	26.8
Indicadores de carencias sociales				
Rezago educativo	42.1	34.0	25.1	14.9
Acceso a los servicios de salud	38.9	34.8	32.6	30.4
Acceso a la seguridad social	92.2	84.0	73.2	49.5
Calidad y espacios de la vivienda	46.1	31.7	18.3	8.7
Acceso a los servicios básicos en la vivienda	83.3	64.3	30.2	8.1
Acceso a la alimentación	43.4	33.4	28.4	20.6
Privación social				
Población con al menos una carencia social	98.4	94.7	85.7	65.7
Población con al menos tres carencias sociales	76.7	59.5	35.7	16.8
Bienestar económico				
Población con un ingreso inferior a la línea de bienestar mínimo	64.7	39.8	22.9	10.9
Población con un ingreso inferior a la línea de bienestar	88.5	75.7	62.0	41.2

Fuente: Estimaciones del CONEVAL con base en el MCS-ENIGH 2010.

4. Indicadores de carencia social

Es esta sección se describe la forma en que se estimaron los indicadores relacionados con las carencias sociales y los métodos directos e indirectos que se utilizaron. Los modelos indirectos se construyeron en cada estrato para cada una de las metodologías implementadas.

4.1 Indicadores de carencia social estimados directamente

La muestra del Censo de Población y Vivienda 2010, como se ha comentado con anterioridad, está diseñada para generar información de cada uno de los municipios del país y contiene información suficiente para estimar de manera directa sólo cuatro de los indicadores de carencia social: rezago educativo, acceso a los servicios de salud, calidad y espacios de la vivienda y acceso a los servicios básicos en la vivienda. Los criterios para determinar estas carencias pueden consultarse en el documento metodológico de la medición de la pobreza (CONEVAL, 2010). Con las definiciones respectivas se construyeron indicadores dicotómicos de cada carencia para todos los individuos en la muestra del Censo de Población y Vivienda 2010.

4.2 Construcción del indicador de carencia de acceso a la seguridad social

La carencia por acceso a la seguridad social es uno de los indicadores para los cuales la muestra del censo no contiene la información suficiente para medirlo directamente, por lo cual se requirió elaborar modelos de estimación. Dado que se contaba con información de los hogares y sus integrantes, así como variables relativas a los municipios a los que pertenecen estos hogares, se

implementaron modelos logísticos jerárquicos con efectos aleatorios en el intercepto,¹⁸ los cuales combinan información de varios niveles de desagregación.

El proceso para estimarlos consistió de tres etapas: 1) selección de las variables determinantes de la carencia de seguridad social; 2) implementación de modelos, y 3) validación de los modelos. Este proceso se siguió para cada uno de los estratos de municipios, por lo que las variables determinantes de la carencia son distintas en cada estrato.

La selección de las variables consistió en identificar un conjunto de factores predictivos de la carencia de seguridad social, para lo cual se revisaron las fuentes de información y la bibliografía sobre seguridad social. Estos factores se enriquecieron con las sugerencias de las y los investigadores académicos que conforman la Comisión Ejecutiva (CE).¹⁹ Además del respaldo teórico para seleccionar las condicionantes a emplear en la estimación, era necesario satisfacer el requisito de que las distribuciones de estos factores en el MCS-ENIGH 2010 y en la muestra del censo fueran similares. Para comparar las distribuciones se realizaron análisis exploratorios con gráficas y comparación de estadísticas básicas y se realizaron pruebas de hipótesis para sustentar que los datos provenían de la misma distribución de probabilidad. Adicionalmente, como parte del proceso de selección de variables, se construyeron regresiones logísticas exploratorias para encontrar algunas de las variables que más explicaban la carencia por acceso a la seguridad social.

El indicador de carencia de seguridad social es un indicador dicotómico: se tiene la carencia o no. Por ello, se utilizaron modelos logísticos jerárquicos de dos niveles (un primer nivel para las variables individuales y familiares, y un segundo nivel para la información municipal). Para cada estrato se estimaron y validaron los modelos que mejor ajustaban la distribución de probabilidad de las variables en cuestión (en este caso, el acceso a la seguridad social). Después, estos modelos fueron aplicados a la información de la muestra del censo.

Los modelos jerárquicos implementados en la muestra del censo generaron la probabilidad de que cada integrante del hogar fuera carente en el indicador de seguridad social. Estas probabilidades fueron convertidas en indicadores dicotómicos. Para decidir si un individuo estaba en situación de carencia en el acceso a la seguridad social, se comparó la probabilidad que le asignó el modelo logístico jerárquico con la media de la probabilidad del municipio al que pertenecía; si la probabilidad del individuo era mayor a la media de la probabilidad municipal entonces el individuo se consideró carente, y no carente en caso contrario. La idea fundamental del método de conversión fue respetar la probabilidad por municipio que el mismo modelo logístico jerárquico estimaba.

Con el procedimiento para estimar el indicador de carencia por el acceso a la seguridad social se tuvo la estimación para cada individuo en la muestra del censo; de esta manera se resolvió el problema de identificación de los individuos con respecto a esa carencia.

¹⁸ Randenbush y Bryk (2002), y Hoffmann (2004)

¹⁹ La Comisión Ejecutiva es el órgano técnico-auxiliar del CONEVAL encargado de ejercer las funciones sustantivas de carácter técnico.

4.3 Construcción del indicador de carencia de acceso a la alimentación

La carencia de acceso a la alimentación es la otra carencia social que no puede ser estimada directamente con la muestra del censo. Este indicador está integrado por una escala compuesta por doce preguntas: la Escala Mexicana de Seguridad Alimentaria (CONEVAL, 2010). Sin embargo, en la muestra asociada al censo únicamente se incluyeron dos de las doce preguntas de la escala alimentaria.

El proceso que se siguió para estimar este indicador fue similar al utilizado para el acceso a la seguridad social. Se seleccionó un conjunto de factores determinantes de la carencia de alimentación en los estratos, se implementaron los modelos, y finalmente, se buscó el mejor ajuste de los modelos con las variables determinantes.

La selección de las variables consistió en identificar los factores condicionantes que tuvieran relación con el acceso a la alimentación en los hogares, para lo cual se revisaron las fuentes de información y la bibliografía sobre el tema. Las variables fueron presentadas a la CE, y de allí se obtuvo una lista de posibles variables determinantes que se exploraron empíricamente.

Para cumplir con los supuestos de los modelos en áreas pequeñas, fue necesario comprobar que las distribuciones de las variables entre las fuentes de información (MCS-ENIGH y muestra del censo) fueran similares. Se realizaron análisis exploratorios con gráficas y comparación de estadísticas básicas y se hicieron pruebas de hipótesis sobre la igualdad de las distribuciones. Otro criterio para seleccionar las variables se realizó mediante regresiones logísticas en cada uno de los estratos tomando como fuente principal el MCS-ENIGH 2010. Al contrastar resultados preliminares obtenidos con los modelos de regresión logística resaltó el hecho de que las estimaciones del MCS-ENIGH en el estado de Guerrero eran muy lejanas a lo que estimaban los modelos. Por lo anterior se decidió estimar cinco modelos, uno correspondiente a cada estrato y, como caso particular, se modeló la carencia de alimentación para el estado de Guerrero.

Una vez definida la partición de los municipios del país en cinco estratos y los factores asociados a la carencia en alimentación, el siguiente paso consistió en implementar la estimación mediante el ajuste de diversos modelos logísticos jerárquicos en el MCS-ENIGH. Una de las pruebas básicas al contrastar resultados fue compararlos con los resultados preliminares de las regresiones logísticas realizadas para determinar las variables independientes de los modelos. Al realizar este contraste resultó que las estimaciones del indicador eran más cercanas para las entidades federativas con los modelos de regresión logística que con los modelos jerárquicos. Esa fue la razón fundamental para implementar modelos de regresión logística para el indicador de carencia en el acceso a la alimentación en la muestra del censo.

Al aplicar los modelos de regresión logística ajustados a la información de la muestra del censo, se estimó la probabilidad de estar en situación de carencia para cada hogar, probabilidad que se imputó a todos los individuos que pertenecen al hogar. Con el propósito de identificar la carencia

de los individuos, se transformó la probabilidad en un indicador dicotómico, con un criterio análogo al aplicado para la carencia de acceso a la seguridad social.²⁰

5. Indicadores relacionados con el espacio del bienestar económico

El ingreso de los hogares es uno de los indicadores de bienestar económico comúnmente utilizado en la medición de la pobreza. Además, es uno de los indicadores establecido en la LGDS, que permite identificar la dimensión económica de la pobreza a partir del análisis de los recursos monetarios de los hogares para la satisfacción de sus necesidades. Esta dimensión se mide con el indicador de ingreso corriente total per cápita (ICTPC).²¹

Para estimar este indicador, el CONEVAL utilizó la información de la muestra del censo y del MCS-ENIGH, ambos para el año 2010. El MCS-ENIGH está diseñado con el propósito de medir con precisión los ingresos de los hogares mexicanos además de las carencias sociales. Por otro lado, la muestra del censo tiene como finalidad profundizar el conocimiento sobre algunas características de las personas a fin de caracterizar en detalle el perfil sociodemográfico de la población. Sin embargo, la información de los ingresos de las personas que reporta la muestra del censo no corresponde al ICTPC, pues solamente considera los ingresos laborales. El objetivo principal era realizar una estimación indirecta del ingreso de los integrantes de los hogares mexicanos que permitiera distinguir si sus ingresos eran superiores o inferiores a la línea de bienestar y a la línea de bienestar mínimo,²² la cual debía ser representativa para todos los municipios del país.

En la sección 2 se detalló que solamente dos de los métodos estudiados, el MPEG y el ELLG, satisfacían todos los criterios generales que se requerían para realizar la medición municipal de la pobreza según lo establece la LGDS. Se demostró que el método ELLG resultó ser el que generó estimaciones más precisas en relación con las cifras estatales.

Por esta razón, la estimación del ingreso se implementó utilizando el paquete de cómputo *PovMap*.²³ En México, el CONEVAL lo había implementado previamente para la estimación de la pobreza por ingresos a escala municipal²⁴ para los años 2000 y 2005. El método ELL permitió combinar la información de la muestra del censo 2010 y la del MCS-ENIGH 2010 para construir una medida de ingreso agregable en todos sus componentes y que, al mismo tiempo, fuera representativa a escala municipal. La ventaja de emplear este método fue que el ICTPC se simuló

²⁰ Para decidir si un individuo se encontraba en situación de carencia en el acceso a la alimentación, se comparó la probabilidad del individuo con la media de la probabilidad del municipio al que pertenece; si la probabilidad del individuo era mayor a la media de la probabilidad municipal entonces el individuo fue considerado como carente, y como no carente en caso contrario.

²¹ Se trata del ingreso del que disponen las personas para la adquisición de bienes y servicios en el mercado, específicamente, el ingreso corriente, el cual representa el flujo de entradas, no necesariamente monetarias (puede incluir los productos disponibles o recibidos en especie, entre otros), que le permiten a los hogares obtener los satisfactores que requieren, sin disminuir los bienes o activos que poseen (CONEVAL 2010).

²² La medición de pobreza utiliza dos líneas de ingreso: la línea de bienestar mínimo, que equivale al valor de la canasta alimentaria por persona al mes; y la línea de bienestar, que equivale al valor total de la canasta alimentaria y de la canasta no alimentaria por persona al mes (CONEVAL, 2010, 2012).

²³ El *PovMap* es un paquete de acceso libre; fue elaborado por el Banco Mundial.

²⁴ La metodología y los resultados de la pobreza por ingresos se pueden consultar en la página electrónica del CONEVAL: www.coneval.gob.mx.

para cada persona en los hogares de la muestra del censo, por lo cual se pudieron realizar agregaciones en distintos niveles (nacional, estatal y municipal) y, al mismo tiempo, fue factible estimar las medidas de desigualdad en la distribución del ingreso, como indicadores del grado de cohesión social de los municipios (índice de Gini, razón de ingresos, promedio del ingreso por municipio).

De manera similar a los modelos relacionados con las carencias sociales, se construyeron modelos para cada estrato. Para ser consistente con los principios que se tomaron en cuenta en las mediciones de pobreza por ingresos de los años 2000 y 2005, se buscaron factores condicionantes del ingreso diferenciados en los ámbitos rural y urbano. Esta decisión hizo necesario dividir cada uno de los cuatro estratos en los ámbitos rural y urbano.²⁵ Sin embargo, para el estrato uno urbano no se contaba con un tamaño de muestra suficiente en el MCS-ENIGH 2010, por lo que se realizaron siete modelos correspondientes a los estratos con su partición urbana y rural: estrato uno rural, estrato dos rural, estrato tres rural, estrato cuatro rural, estrato uno y dos urbano, estrato tres urbano y estrato cuatro urbano (consultar la estimación de los indicadores relacionados con el ingreso según estrato rural y urbano en el cuadro C.8 del anexo C).

Con este método se obtuvo la distribución de las variables relacionadas con el ingreso (factores condicionantes del ingreso) tanto para la encuesta como para la muestra del censo. Con la finalidad de definir el tipo de variables que serían seleccionadas, se revisó el contenido temático del cuestionario de la encuesta y de la muestra del censo, lo que permitió seleccionar un conjunto de variables demográficas y educativas, así como algunos factores ocupacionales y laborales, que podían ser consideradas conceptual y operativamente comparables y relevantes. Cabe mencionar que también se añadieron variables predictivas a nivel municipal, las cuales sirvieron para mejorar la bondad de ajuste de los modelos.

Con el programa *PovMap* se realizaron pruebas de los diferentes modelos por estrato y ámbito urbano y rural para verificar la correlación de cada una de las covariables (variables determinantes) con el ingreso. Los modelos se validaron y se escogieron aquellos con el mayor coeficiente de determinación.

El método ELL construye la distribución del logaritmo del ICTPC condicional a las variables determinantes a través de simulaciones generadas con la técnica de remuestreo (*bootstrapping* o muestreo-remuestreo, Gutiérrez-Peña, 1997). Este procedimiento se realizó para obtener cien réplicas para cada hogar en la muestra del censo. De esta manera, se puede calcular la distribución condicional de los indicadores relacionados con el bienestar económico en cada municipio, y, al mismo tiempo, estimar los errores cuadráticos medios respectivos.

El proceso para calcular los indicadores, una vez que se contó con cien simulaciones del logaritmo del ICTPC para cada hogar fue el siguiente: se contrastó el ingreso con la línea de bienestar y con la línea de bienestar mínimo, con lo cual se imputaron para cada réplica dos variables dicotómicas. La primera variable dicotómica se construyó de la siguiente forma: se le asignó uno si el ingreso era inferior a la línea de bienestar, o cero si el ingreso era mayor o igual. De la misma forma se construyó una segunda variable dicotómica definida como uno si el ingreso era inferior a la línea de

²⁵ Una localidad se considera rural si su población es menor a 2,500 habitantes, (CONEVAL, 2010).

bienestar mínimo y cero si el ingreso era mayor o igual. Se debe notar que, como se ha explicado, se tienen estas variables dicotómicas para cada una de las cien replicas de todos los hogares de la muestra del censo.

El porcentaje de la población con ingreso inferior a la línea de bienestar se obtuvo en dos pasos. Primero, en cada simulación se calculó el promedio de la variable dicotómica que indica si el ingreso del hogar es inferior a la línea de bienestar, ponderado por el factor de expansión. Segundo, el estimador de la incidencia de la pobreza por ingresos se estima a través del promedio de la incidencia de cada simulación (consultar la descripción matemática en el anexo A). De manera similar se calculó el indicador para el porcentaje de la población con ingreso inferior a la línea de bienestar mínimo.

6. Integración de los indicadores de pobreza (ELLG) y calibración de la población

En la presente sección se tratarán dos aspectos de la medición multidimensional de la pobreza a escala municipal. En primer lugar, se abordará la forma en que se integraron los indicadores estimados en las secciones anteriores para conformar los cuadrantes de pobreza (figura A.1 del anexo A; y CONEVAL, 2010), conformados por la población en situación de pobreza (desagregada en pobreza extrema y pobreza moderada), vulnerabilidad por carencias sociales, vulnerabilidad por ingresos y población no pobre y no vulnerable. En segundo lugar, se tratará el proceso de calibración de las estimaciones por entidad federativa y por municipios.

Recordemos que los tres espacios asociados a la medición de la pobreza son el bienestar económico, las carencias sociales y el contexto territorial. El indicador asociado al bienestar económico es el ICTPC; los indicadores dicotómicos de las seis carencias sociales son el rezago educativo, el acceso a los servicios de salud, el acceso a la seguridad social, la calidad y espacios de la vivienda, el acceso a los servicios básicos en la vivienda y el acceso a la alimentación. Los indicadores de cohesión social son la razón de ingresos (ingreso medio de la población no pobre y no vulnerable entre el ingreso medio de la población pobre extrema) y el coeficiente de Gini. Además, y de manera independiente, se estimó el índice de polarización social de los municipios, según se establece en la metodología para la medición multidimensional de la pobreza en México (CONEVAL, 2010).

6.1 Integración de los indicadores de pobreza (metodología ELLG)

La construcción de los cuadrantes de pobreza requirió del espacio de bienestar económico (ingreso) y de las carencias sociales. Con los seis indicadores de carencia social se construyó el Índice de Privación Social (IPS), el cual se obtiene al sumar las carencias sociales de cada individuo. Para el ingreso se contaba con cien simulaciones del ingreso promedio de los hogares y también cien simulaciones de las dos variables dicotómicas relacionadas con la población con ingreso inferior a la línea de bienestar y con ingreso inferior a la línea de bienestar mínimo, obtenidas con el método ELL.

El desafío radicaba en combinar las cien simulaciones de las variables dicotómicas relacionadas con el ingreso con el índice de privación social para construir los cuadrantes de la pobreza. Este problema se resolvió al generalizar el método ELL para la medición multidimensional de la pobreza, método al que hemos denominado ELLG, para cada una de las variables de los cuadrantes de pobreza. Dado que se tenía el IPS, éste se pudo contrastar con cada una de las simulaciones relacionadas con el ingreso y generar cien simulaciones de pobreza como indicador dicotómico y cien simulaciones de pobreza extrema. Al ser ambas funciones del ingreso cumplían con los supuestos del método ELL de ser funciones de probabilidad condicionales a las variables determinantes. De esta forma, el estimador de pobreza para cada municipio se obtuvo en dos pasos: primero, se calculó el promedio de las simulaciones de pobreza, ponderado por el factor de expansión para cada municipio en cada una de las simulaciones; segundo, el estimador de incidencia de la pobreza resulta de calcular la media de las incidencias de pobreza en las distintas simulaciones.

El IPS es una estimación;²⁶ sin embargo, al ser una de las dos variables determinantes de los cuadrantes de pobreza (por lo tanto de la pobreza y pobreza extrema), para efectos de este ejercicio de medición multidimensional de la pobreza (ELLG) se consideró como una variable dada. Nótese que este procedimiento no afecta las estimaciones puntuales de pobreza y sus componentes, pero sí el cálculo de los errores cuadráticos medios.

El resto de los indicadores se obtiene por diferencias. Así, la incidencia de la pobreza moderada de un municipio es la diferencia entre la incidencia de pobreza menos la incidencia de pobreza extrema. El porcentaje de población que es vulnerable por carencias sociales se obtiene como la diferencia del porcentaje de población con al menos una carencia social menos el porcentaje de población en situación de pobreza. El porcentaje de población que es vulnerable por ingresos se estima como la diferencia entre el porcentaje de población con ingresos menores a la línea de bienestar y el porcentaje de población pobre. Finalmente, la población no pobre y no vulnerable se determina como el complemento de la población pobre o vulnerable, ya sea por carencias o por ingresos.

6.2 Calibración de las estimaciones de pobreza

Una vez estimada la pobreza y sus indicadores desagregados para cada uno de los municipios del país, se tuvo que resolver el problema de la consistencia entre las estimaciones de los municipios y las entidades federativas a las que pertenecen. Las estimaciones por entidad federativa tuvieron como única fuente el MCS-ENIGH 2010, mientras que las estimaciones municipales derivaron en parte de variables recabadas en la muestra del censo y en parte de imputar los indicadores que no contenía. Se tenían las posibilidades de calibrar a la información del MCS-ENIGH o a la información de la muestra del censo por contener un mayor tamaño de muestra. Se decidió calibrar al MCS-ENIGH 2010 por las siguientes razones:

²⁶ Como se comentó al inicio de la sección 6.1 el IPS es la suma de los seis indicadores dicotómicos asociados a las carencias sociales, los cuales son estimados de manera directa (sección 4.1) y de forma indirecta (secciones 4.2 y 4.3), por lo que el IPS en realidad también es una variable aleatoria.

1. Porque el MCS-ENIGH es una muestra diseñada específicamente para calcular todos los indicadores de pobreza, con lo cual se estiman directamente a nivel entidad federativa y nacional, mientras que en la muestra del censo algunas estimaciones se obtienen de forma indirecta a través de modelos estadísticos.
2. Porque la serie de tiempo de pobreza nacional y por entidad federativa se genera con una periodicidad bienal. No se deseaba introducir una variación en la serie de tiempo que pudiera atribuirse a variaciones de la fuente de información, del método de estimación o como resultado de la calibración.
3. Porque las estimaciones por entidad federativa habían sido dadas a conocer previamente a las estimaciones por municipios y, por tanto, ya se encontraban posicionadas en la opinión pública.

Para abordar este problema se utilizó el método de calibración *logit* (método de calibración logístico)²⁷ descrito en Deville y Särndal (1992), en Deville y colaboradores (1993) y en Vargas y colaboradores (2011, Calibración de indicadores socioeconómicos).

Con el método de calibración logístico se construye una función multivariada. Una vez obtenidos los totales por entidad federativa de cada una de las variables que se desea calibrar, se plantea un sistema de ecuaciones cuyo propósito es encontrar nuevos factores de expansión “calibrados”. Las variables utilizadas para calibrar fueron los totales de los indicadores básicos de pobreza: el rezago educativo, el acceso a los servicios de salud, la calidad y espacios de la vivienda, el acceso a los servicios básicos en la vivienda, el acceso a la alimentación, el acceso a la seguridad social, la población con ingreso inferior a la línea de bienestar y de bienestar mínimo, la población con al menos una carencia social, la población con tres o más carencias sociales, la pobreza, la pobreza extrema y la población de las entidades federativas. Con este método se obtuvieron los factores calibrados, los cuales tuvieron como propósito que las estimaciones agregadas de los municipios coincidieran con las de la entidad federativa de la cual forman parte.

El método de calibración logístico tiene la ventaja de que los factores calibrados son función de los factores originales y se acotan a un intervalo definido, con lo cual el método respeta lo más posible los pesos originales. Con los factores calibrados se recalcularon los indicadores municipales a partir del procedimiento explicado en la primera parte de la presente sección. Cabe mencionar que se revisó que el ordenamiento de los municipios en pobreza y pobreza extrema no tuviera cambios grandes respecto de las estimaciones previas a la calibración.

7. Profundidad e intensidad de la pobreza

La profundidad asociada al índice de privación social se reporta mediante la proporción promedio de carencias sociales. Este indicador se estimó para los siguientes grupos: la población en pobreza, en pobreza extrema, en pobreza moderada y para la población vulnerable por carencias sociales (CONEVAL, 2010).

²⁷ En el documento se llamará método de calibración logístico, por su traducción al español.

Una vez obtenidas las medidas de incidencia de la pobreza, se calculó la proporción de carencias que, en promedio, experimentaban las personas en los grupos asociados a la medición de pobreza mencionados. Los indicadores de profundidad se construyen al dividir el índice de privación social de las personas que están en situación de pobreza que se está calculando entre el total de indicadores asociados a las seis carencias sociales. En seguida, se calcula el promedio, ponderado con los factores de expansión calibrados, del índice de privación social, restringido al grupo de interés (cuadrante).

Las medidas de intensidad se construyen a partir de la multiplicación de la medida de incidencia y la medida de profundidad (CONEVAL, 2010). Estas medidas permiten realizar diagnósticos sensibles a los cambios en las condiciones de vida de la población en situación de pobreza (Alkire y Foster, 2007).

8. Indicadores de cohesión social

Para cada municipio se reportaron los siguientes indicadores de cohesión social a nivel municipal: el índice de Gini, la razón de ingreso y el grado de polarización (CONEVAL, 2010).

El índice de Gini es una medida de la desigualdad en la distribución del ingreso. Este índice varía entre cero y uno. Mientras más alto es su valor mayor desigualdad existe en la distribución del ingreso (Deaton, 1997 y Ray, 1998). El índice de Gini se calculó con el ICTPC de los hogares, estimado según el método ELL. Las cien simulaciones de ingreso originaron cien índices de Gini para cada municipio. El promedio de las simulaciones fue el estimador del coeficiente de Gini municipal.

Del mismo modo que el coeficiente de Gini, las razones del ingreso se estimaron por el método ELL. Estas razones muestran el cociente del ingreso medio de la población no pobre y no vulnerable entre el ingreso medio de la población en situación de pobreza extrema. Estas razones indican cuántas veces es mayor el ingreso de la población no pobre y no vulnerable con respecto a la población en pobreza extrema. Ambas medidas se utilizan como indicadores de la desigualdad del ingreso en los municipios. Se supone que los municipios con mayor desigualdad tienen un menor grado de cohesión social.

El grado de polarización social de los municipios se calculó con información del Índice de Marginación por localidad para 2010 publicado por el Consejo Nacional de Población (CONAPO, 2012).²⁸

Una vez que se clasificó a los municipios según la polarización social, se definió el grado de cohesión social. El municipio se consideró de baja cohesión social si el municipio era polarizado o con polo izquierdo. Por el contrario, era de alta cohesión social si se trataba de un municipio de polo derecho o sin polo.

²⁸ La descripción detallada de la forma en que se construye este indicador puede consultarse en CONEVAL, 2010.

Conclusiones

De acuerdo con los ordenamientos de la LGDS, el CONEVAL debe medir la pobreza municipal cada cinco años. En diciembre de 2011, en cumplimiento con el mandato de la ley, se presentaron los resultados de la primera medición multidimensional de pobreza a escala municipal. Este documento metodológico expone las principales decisiones que se tomaron para llevar a cabo la medición municipal de la pobreza. De esta manera se cumple con una de las normas básicas que rigen el quehacer institucional del CONEVAL, es decir, utilizar procedimientos objetivos, rigurosos y transparentes.

Al hacer públicos los métodos empleados por el CONEVAL, se promueve el intercambio de observaciones, críticas y comentarios encaminados a mejorar el estudio y medición de la pobreza a escala municipal. Asimismo, la presentación detallada de los procesos llevados a cabo permite apreciar las diversas opciones con las que se contaron y los caminos que se eligieron, lo que permite apreciar sus fortalezas, debilidades y necesidades de investigación en el futuro.

Al inicio del documento se describe la problemática central de la medición multidimensional de la pobreza a escala municipal en México. En efecto, a diferencia de la medición nacional y estatal, no se contaba con una fuente de información que permitiera la estimación directa de la pobreza multidimensional en los municipios. Ante esta situación, fue necesario explorar, analizar, implementar y seleccionar un grupo de metodologías que permitieran estimar indirectamente algunos de los indicadores de pobreza. El proceso de investigación, que tuvo una duración de tres años, involucró la participación de diversas instituciones e investigadores expertos en la materia.

Con este documento metodológico, el CONEVAL no solo impulsa el estudio y comprensión de la medición multidimensional de la pobreza, sino que contribuye a la generación de información que permite evaluar de mejor manera la política social mexicana. Con la información detallada de la pobreza a escala municipal, es posible identificar las necesidades específicas de grupos poblacionales y diseñar programas sociales que se orienten de mejor manera a las poblaciones más necesitadas.

Este documento debe considerarse como un primer paso en el análisis de los rezagos sociales que persisten en la sociedad mexicana. Los requerimientos y las limitaciones de la información establecen retos considerables, entre los cuales destacan los siguientes: contar con información con mayor desagregación para instrumentar de manera más eficiente las políticas públicas; mejorar los indicadores asociados a la pobreza; identificar regiones en el país que nos permitan contextualizar el fenómeno de la pobreza. La contribución fundamental de este trabajo de estimación es que ahora se dispone de información desagregada de las múltiples dimensiones de la pobreza a nivel municipal, y no solamente del ingreso, como se estima usualmente la pobreza por áreas pequeñas en la bibliografía especializada.

Desde el punto de vista metodológico, se requiere reforzar los procedimientos estadísticos para medir multidimensionalmente la pobreza y sus dimensiones, lo que exige continuar explorando las distintas metodologías existentes, valorar sus fortalezas y debilidades, la precisión estadística de los métodos de estimación y la consistencia de sus resultados. Una tarea particularmente relevante

atañe a la estimación simultánea, y a nivel individual, de los tres indicadores que no pueden obtenerse directamente de la muestra del censo de población y vivienda.

Referencias bibliográficas

Alkire, S y Foster, J. (2007). "Counting and Multidimensional Poverty Measurement", *OPHI Working Paper Series*, Oxford Poverty & Human Development Initiative (OPHI), Oxford, (OPHI Working Paper, 7) recuperado de http://www.ophi.org.uk/pubs/Alkire_Foster_CountingMultidimensionalPoverty.pdf (2008, 18 de junio).

Aragon, Y., Casanova, S., Chambers, R. y Leconte, E. (2005). "Conditional Ordering Using Nonparametric Expectiles". *Journal of Official Statistics*, 21. (pp. 617-633).

Banerjee, S., Gelfand, A. E. y Carlin, B. P. (2003). *Hierarchical Modeling and Analysis for Spatial Data*. New York: Chapman & Hall/CRC.

Battese, G.E., Harter, R.M., y Fuller, W.A. (1988). "An Error Component Model for Prediction of County Crop Areas Using Survey and Satellite Data". *Journal of the American Statistical Association*, 83. (pp. 28-36).

Bernardo, J. M. y Smith, A. F. M. (2000). *Bayesian Theory*. New York: Wiley.

Breckling J. y Chambers, R. (1988). "M-quantiles". *Biometrika*, 75, (pp. 761-771).

Bowman, A.W., Hall, P. y Prvan, T. (1998). "Bandwidth Selection for the Smoothing of Distribution Functions". *Biometrika*, 85, (pp. 799-808).

Brunsdon, C.F., Fotheringham, A.S. y Charlton, M. (1996). "Geographically Weighted Regression - A Method for Exploring Spatial Non-Stationarity." *Geographical Analysis*, 28. (pp. 281-298).

Chambers, R. (1986). "Outlier Robust Finite Population Estimation". *Journal of the American Statistical Association*, 81. (pp. 1063-1069).

Chambers, R., Chandra, H. y Tzavidis, N. (2010). "On Robust Mean Squared Error Estimation for Linear Predictors for Domains". *The University of Manchester, Cathie Marsh Centre for Census and Survey Research, CCSR Working Paper 2007-10*.

Chambers, R., Dorfman, A.H. y Hall, P. (1992). "Properties of Estimators of the Finite Population Distribution Function". *Biometrika*, 79. (pp. 577-582).

Chambers, R., Dorfman, A.H. y Wehrly, T.E. (1993). "Bias Robust Estimation in Finite Populations Using Nonparametric Calibration ". *Journal of the American Statistical Association*, 88.(pp. 268-277).

Chambers, R. y Dunstan, R. (1986). "Estimating Distribution Functions from Survey Data". *Biometrika*, 73. (pp. 597-604).

Chambers, R. y Tzavidis, N. (2006). "M-quantile Models for Small Area Estimation". *Biometrika*, 93. (pp. 255-268).

Chantala, K. (2001). *Using STATA to Analyze Data from a Sample Survey*. UNC Chapel Hill. Carolina Population Center.

CONAPO (2012). *Índice de marginación por localidad 2010*. México, CONAPO.

CONEVAL (2007). *Los mapas de pobreza en México. Anexo técnico metodológico*. México. CONEVAL.

CONEVAL (2012). *Construcción de las Líneas de Bienestar. Documento Metodológico*. México.

____ (2010). *Metodología para la medición multidimensional de la pobreza en México*. México. CONEVAL.

Datta, G.S., Fay, R.E. y Ghosh, M. (1991). "Hierarchical and Empirical Bayes Multivariate Analysis in Small Area Estimation". *Proceedings of Bureau of the Census 1991 Annual Research Conference*. U.S. Bureau of the Census. Washington, DC. (pp. 63-79).

Deaton, A. (1997). *The Analysis of Household Surveys: A Microeconometric Approach to Development Policy*. The World Bank, The Johns Hopkins University Press, Baltimore, pp. 148-154.

Demombynes, G., Elbers, C., Lanjouw, J. O., Lanjouw, P. (2008). "How Good a Map? Putting Small Area Estimation to the Test". *Rivista Internazionale di Scienze Sociali*, Vol. 4, pp. 465-493.

Deville, J. C. y C. E. Särndal (1992). "Calibration Estimators in Survey Sampling". *Journal of the American Statistical Association*, (Vol. 87), No. 418. (pp. 376-382).

Deville, J.-C., Särndal, C. E. y Sautory, O. (1993). "Generalized Raking Procedures in Survey Sampling". *Journal of the American Statistical Association*, Vol. 88, No. 423, pp. 1013-1020.

Diario Oficial de la Federación (16 de junio de 2010). *Lineamientos y criterios generales para la definición, identificación y medición de la pobreza*, México, recuperado de http://dof.gob.mx/nota_detalle.php?codigo=5146940&fecha=16/06/2010 (2011, 22 de noviembre).

Dijkstra, E. W. (1959). "A Note on Two Problems in Connexion with Graphs". *Numerische Mathematik*, 1. (pp. 269-271).

Duan, N. (1983). "Smearing Estimate: A Nonparametric Retransformation Method". *Journal of the American Statistical Association*, 78. (pp. 605-610).

Elbers, C., Lanjouw, J. O. y Lanjouw, P. (2002). "Micro-level Estimation of Welfare". *Policy Research Working Paper 2911*, The World Bank, Washington, DC.

Elbers, C., Lanjouw, J. O. y Lanjouw, P. (2003). "Micro-level Estimation of Poverty and Inequality". *Econometrica*, (Vol. 71), No. 1. (pp. 355-364).

Fay, R.E. y Herriot, R.A. (1979). "Estimation of Income from Small Places: An Application of James-Stein Procedures to Census Data". *Journal of the American Statistical Association*, 74. (pp. 269-277).

Ghosh, M., Natarajan, K., Stroud, T.W.F., y Carlin, B.P. (1998). "Generalized Linear Models for Small Area Estimation". *Journal of the American Statistical Association*, 93. (pp. 273-281).

Ghosh, M., Natarajan, K., Waller, L.A., y Kim, D. (1999). "Hierarchical Bayes GLMs for the Analysis of Spatial Data: An Application to Disease Mapping". *Journal of Statistical Planning and Inference*, 75. (pp. 305-318).

Ghosh, M. y Rao, J.N.K. (1994). "Small Area Estimation: An Appraisal (with discussion)". *Statistical Science*, 9. (pp. 65-93).

González-Manteiga, W., Lombardía, M.J., Molina, I., Morales, D., y Santamaría, L. (2008). "Bootstrap Mean Square Error of a Small-Area EBLUP". *Journal of Statistical Computation and Simulation*, 78. (pp. 443-462).

Gould, W. W., J. S. Pitblado, and B. P. Poi. (2010). *Maximum Likelihood Estimation with Stata*, Estados Unidos, College Station, Stata Press.

Gutiérrez-Peña, E. (1997). *Métodos computacionales en la inferencia bayesiana*. Monografía IIMAS-UNAM, (Vol. 6), No. 15.

Hagenaars, Jacques y Allan McCutcheon (2002). *Applied Latent Class Analysis*, Reino Unido, Cambridge University Press.

Hoffmann, J. P. (2004). *Generalized Linear Models and applied approach*. Estados Unidos, Pearson Education Inc.

INEGI (2000). *ENIGH-2000, Encuesta Nacional de Ingresos y Gastos de los Hogares, documento metodológico*. INEGI, recuperado de <http://www.inegi.org.mx> (2012, 23 de agosto).

Kingman, JFC. (1993). *Poisson Processes*. Oxford University Press.

Kokic, P., Chambers, R., Breckling, J. y Beare, S. (1997). "A Measure of Production Performance". *Journal of Business and Economic Statistics*, 15. (pp. 445-451).

Koenker, R. (2005). *Quantile Regression*. New York: Cambridge University Press

Koenker, R. y Bassett, G. (1978). "Regression Quantiles". *Econometrica*, 46. (pp. 33-50).

Koenker, R. y D'OREY, V. (1987). "Computing Regression Quantiles". *Journal of the Royal Statistical Society Series C*, 36. (pp. 383-393).

Li, Q. and Racine, J.S. (2007). *Nonparametric Econometrics: Theory and Practice*. Princeton: Princeton University Press.

Lombardia M.J., Gonzalez-Manteiga, W. y Prada-Sanchez, J.M. (2003). "Bootstrapping the Dorfman-Hall-Chambers-Dunstan Estimator of a Finite Population Distribution Function". *Journal of Nonparametric Statistics*, 16,.(pp. 63-90).

Mardia, K.V., Kent, J.T. y Bibby, J.M. (1997). *Multivariate Analysis (Probability and Mathematical statistics)*. Reino Unido, Academic Press

- Molina, I. y Rao, J.N.K. (2010). "Small Area Estimation of Poverty Indicators". *To appear in the Canadian Journal of Statistics*.
- Muthén, L. y Muthén B. (2010). *Mplus. Statistical Analysis with Latent Variables. User's Guide*, Estados Unidos, Muthén & Muthén.
- Nieto, L. E. (2011). "Método Bayesiano". Delfino Vargas y colaboradores. *Metodología de ajustes e imputación de indicadores de la pobreza por ingreso en áreas pequeñas*. México. El Colegio de México y CONEVAL, mimeo.
- Pfefferman, D. (2002). "Small Area Estimation. New Developments and Directions". *International Statistical Review*, 79. (pp. 125-143).
- Prasad, N.G.N y Rao, J.N.K. (1990). "The Estimation of the Mean Squared Error of Small Area Estimators". *Journal of the American Statistical Association*, 85. (pp. 163-171).
- Pratesi, M., Tzavidis, N., Molina, I., Salvati, N., Marchetti, S., Giusti, C. y Domingo, M. (2010). "Recent Advances in Model-based Small Area Estimation of Poverty Indicators". Working Paper.
- Rao, J. N. K. (2003). *Small Area Estimation*, Estados Unidos, John Wiley & Sons, Inc.
- Rao, J.N.K., Kovar, J.G. y Mantel, H.J. (1990). "On Estimating Distribution Functions and Quantiles from Survey Data Using Auxiliary Information". *Biometrika*, 77. (pp. 365-375).
- Raudenbusch, S. W. y Anthony S. Bryk (2002). *Hierarchical Linear Models. Applications and Data Analysis Methods. Second Edition*. Estados Unidos. Sage Publications, Inc.
- Raudenbusch, S. W., Bryk, A., Congdon, R. (2011). *HLM: Hierarchical Linear and Nonlinear Modeling*. Estados Unidos, Scientific Software International, Inc.
- Ray, D. (1998). *Development Economics*. Estados Unidos, Princeton University Press.
- Raydan, M. (1997). *The Barzilai and Borwein Gradient Method for the Large Scale Unconstrained Minimization Problem*. SIAM Journal of Optimization, Vol. 7, pp. 26-33.
- Richardson, A.M. (1997). "Bounded Influence Estimation in the Mixed Linear Model". *Journal of the American Statistical Association*, 92. (pp. 154-161).
- Richardson, A.M. y Welsh, A.H. (1995). "Robust Estimation in the Mixed Linear Model". *Biometrics*, 51. (pp. 1429-1439).
- Rowe, D. B. (2003). *Multivariate Bayesian Statistics*. Ney York. Chapman and Hall.
- Royall, R.M. y Cumberland, W.G. (1978). "Variance Estimation in Finite Population Sampling". *Journal of the American Statistical Association*, 73. (pp. 351-358).
- Särndal, C. E., B. Swensson y J. Wretman. (2003). *Model Assisted Survey Sampling*. Estados Unidos. Springer-Verlag.

- Salvati, N., Tzavidis, N., Chambers, R. y Pratesi, M. (2010). "M-quantile Geographically Weighted Regression for Small Area Estimation". [Invited to review and resubmit].
- Sarndal, C. E., Swenson, B. y Wretman, J. (1997). *Model Assisted Survey Sampling*. Springer.
- Sharon, L. (1999). *Muestreo Diseño y Análisis*. International Thomson Editores.
- Spector, P. (2008). *Data Manipulation with R*. Estados Unidos. Springer Science+Business Media, LLC.
- Stata Press. (2011). *Stata Programming Reference Manual. Release 12*. Estados Unidos. Stata Press.
- Tarozzi, A. (2011). "Can Census Data Alone Signal Heterogeneity in the Estimation of Poverty Maps?". *Journal of Development Economics*. 95. 170-185.
- Tarozzi A. Deaton A. (2009). "Using Census and Survey Data to Estimate Poverty and Inequality for Small Areas". *Review of Economics and Statistics*. 91. 773-779.
- Varadhan, R. y Gilbert, P. D. (2009). *BB: An R Package for Solving a Large System of Nonlinear Equations and for Optimizing a High-Dimensional Nonlinear Objective Function*. J. Statistical Software, Vol. 32, No. 4.
- Vargas, D., Méndez, I., Moreno, H., Nieto, L. E., Brunson, C., Tzavidis, N., Kwang-Kim, J. y Pérez, G. (2010). *Metodología de ajustes e imputación de indicadores de la pobreza por ingreso en áreas pequeñas*, México, El Colegio de México y CONEVAL, mimeo.
- Vargas, D., Pérez, G., De la Vega, J., Carmona, C. y Luján, C. (2011). *Calibración de indicadores socioeconómicos*. México, El Colegio de México y CONEVAL, mimeo.
- Vilma, N. *Small Area Estimation in Practice*. Vilnius Gediminas Technical University. Statistics Lithuania, Lithuania e-mail: nekrasaite.vilma@gmail.com
- Wang, S. y Dorfman, A.H. (1996). "A New Estimator of the Finite Population Distribution Function". *Biometrika*, 83, (pp. 639-652).
- Welsh, A.H. y Ronchetti, E. (1998). "Bias-calibrated Estimation from Sample Surveys Containing Outliers". *Journal of the Royal Statistical Society Series B*, 60. (pp. 413-428).
- You, Y. y Rao, J.N.K. (2002). "A Pseudo-empirical Best Linear Unbiased Prediction Approach to Small Area Estimation Using Survey Weights". *Canadian Journal of Statistics*, 30. (pp. 431-439).

A. Anexo matemático

A.1 Resumen

En este anexo se describen las expresiones matemáticas correspondientes a los modelos estadísticos usados para estimar la pobreza a nivel municipal. Las definiciones de los indicadores se tomaron tal como están señaladas en la metodología para la medición multidimensional de la pobreza en México (CONEVAL, 2010). En la sección A.1.1 y A.1.2 se describen los parámetros de interés así como la notación que se utilizó. En la sección dos se describe brevemente el modelo propuesto en Elbers y colaboradores (2002 y 2003) para la estimación del indicador del ingreso; este modelo es conocido como modelo ELL. Con el modelo ELL se generan replicas completas de la muestra del censo con la variable ingreso estimada para cada hogar e individuo. A partir de estas réplicas es posible calcular fácilmente casi cualquier medida basada en el ingreso incluyendo el coeficiente de Gini y las razones de ingreso, indicadores del grado de cohesión social (Tarozzi y Deaton, 2009). En la sección tres se describe como se adopta el modelo de Elbers y colaboradores generalizado a la medición multidimensional de la pobreza (ELLG) para la estimación de los indicadores que corresponden a los cuadrantes de pobreza (figura A.1).

Cabe mencionar que en la sección tres se aplica la metodología ELLG suponiendo que el ingreso es la única variable que se debe estimar. Sin embargo, el índice de privación social (IPS) construido a partir de seis indicadores asociados a carencias sociales también debe ser estimado, a partir de la agregación de los indicadores de carencias sociales. En particular el indicador de carencia por acceso a la seguridad social y el indicador de carencia por acceso a la alimentación, fueron estimados con modelos estadísticos (sección A.4). Por ello, es preciso reconocer que las funciones de distribución de las variables asociadas a los cuadrantes de pobreza son condicionales tanto a los factores explicativos del ingreso como al IPS. Además, en esa sección, cuando se describe la metodología, se incluye una etapa en la cual se calibran los factores de expansión de la muestra del censo a partir de los resultados estatales obtenidos con el MCS-ENIGH (sección A.6). Este proceso de calibración no forma parte de la propuesta original de la metodología ELLG; sin embargo, es importante en el contexto de México, pues las estimaciones municipales (de áreas pequeñas) debían ser consistentes con las estimaciones estatales de pobreza que ya habían sido presentadas previamente por el CONEVAL.

Aun cuando los municipios están agrupados en estratos, la mayor parte de la metodología se aplica de manera independiente en cada estrato y por lo tanto, la descripción de los modelos se realiza, sin pérdida de generalidad, sin incluir un subíndice que haga referencia explícita a los estratos. Es importante mencionar que el conjunto de variables independientes o predictores usados en los modelos varía de estrato a estrato. En la sección tres, cuando la metodología se aplica de manera global, es decir, para todos los estratos al mismo tiempo, se indica que así se está realizando.

En la sección cuatro se describen los modelos de regresión logística usados para la estimación de los indicadores de carencia por acceso a la seguridad social y de carencia por acceso a la alimentación. Posteriormente, en la sección cinco, se describe la construcción del índice de

privación social. Finalmente, en la sección seis se puntualiza el método de calibración que se implementó.

A.1.1 Identificación y notación básica

En esta sección se describen los parámetros de interés que se estiman para la identificación de la pobreza a nivel municipal. Considérese una población finita de individuos de tamaño N que está particionada en D áreas pequeñas. En este caso, las áreas pequeñas son los municipios. Cada área pequeña tiene tamaño $N_1, \dots, N_D \Rightarrow \sum_{d=1}^D N_d = N$.

La metodología de medición de la pobreza señala que la definición de pobreza debe tomar en cuenta dos espacios: el del bienestar económico y el de los derechos sociales. Para medir estos espacios, el CONEVAL considera el ingreso y el índice de privación social (IPS) y los combina para identificar a la población en situación de pobreza.

Sea $y_{dj} \geq 0$, el ingreso (para el presente ejercicio el ICTPC) del individuo j en el municipio d y sea $0 \leq IPS_{dj} \leq 6$, el índice de privación social del individuo j en el municipio d . Asimismo, sea $I(A)$ la función indicadora tal que si el evento A es cierto, entonces, $I(A) = 1$, y en caso contrario, $I(A) = 0$.

Cada individuo se clasifica en uno y sólo uno de los cinco cuadrantes siguientes: población en situación de pobreza moderada; población en situación de pobreza extrema; población vulnerable por carencias sociales; población vulnerable por ingresos, y población no pobre y no vulnerable. Para lo anterior, se definen las variables dicotómicas $C^{(1)}, C^{(2)}, C^{(3)}, C^{(4)}$ y $C^{(5)}$, correspondientes a cada uno de estos cuadrantes (figura A.1).

Sean $lb_{min_{dj}}$ y lb_{dj} las líneas de bienestar mínimo y de bienestar para el individuo dj , respectivamente. Se tiene que $lb_{min_{dj}} < lb_{dj}$. Además ambos valores dependen solamente de si la persona j reside en una localidad rural o urbana.²⁹ Sean $C^{(k)}$, $k = 1, \dots, 6$, las variables dicotómicas que definen la situación de cada individuo j de la población de acuerdo con la siguiente formulación:

- Persona en situación de pobreza moderada, $C^{(1)}$.

$$C_{dj}^{(1)} = I \left[\left(lb_{min_{dj}} \leq y_{dj} < lb_{dj} \text{ y } 0 < IPS_{dj} \right) \cup \left(y_{dj} < lb_{min_{dj}} \text{ y } 0 < IPS_{dj} < 3 \right) \right].$$

A.1

- Persona en situación de pobreza extrema, $C^{(2)}$.

$$C_{dj}^{(2)} = I \left(y_{dj} < lb_{min_{dj}} \text{ y } 3 \leq IPS_{dj} \right).$$

A.2

²⁹ Una localidad es rural si tiene menos de 2,500 habitantes (CONEVAL, 2010)

- Persona vulnerable por carencias sociales, $C^{(3)}$.

$$C_{dj}^{(3)} = I(lb_{dj} \leq y_{dj} \text{ y } 0 < IPS_{dj}).$$

A.3

- Persona vulnerable por ingresos, $C^{(4)}$.

$$C_{dj}^{(4)} = I(y_{dj} < lb_{dj} \text{ y } IPS_{dj} = 0).$$

A.4

- Persona no pobre y no vulnerable, $C^{(5)}$.

$$C_{dj}^{(5)} = I(lb_{dj} \leq y_{dj} \text{ y } IPS_{dj} = 0).$$

A.5

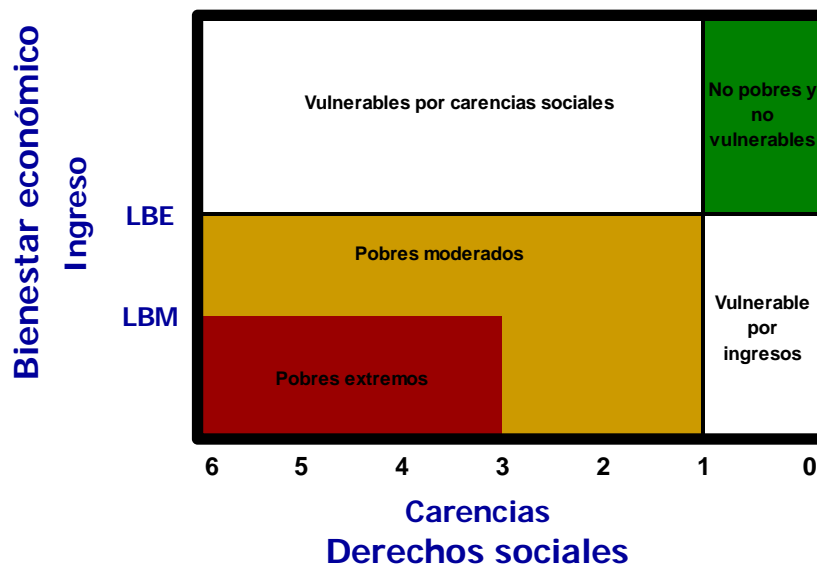
Se dice que un individuo pertenece a la población en situación de pobreza si es parte de la población en situación de pobreza moderada o bien a la población en situación de pobreza extrema.³⁰ Esta situación se determina formalmente mediante la expresión en A.6.

- Persona en situación de pobreza, $C^{(6)}$.

$$C_{dj}^{(6)} = C_{dj}^{(1)} + C_{dj}^{(2)}.$$

A.6

Figura A.1
Cuadrantes de la pobreza



Fuente: Metodología para la medición multidimensional de la pobreza en México (CONEVAL, 2010).
Nota: La población pobre es la población pobre extrema o pobre moderada.

³⁰ La pobreza moderada se deriva de la diferencia entre la pobreza y la pobreza extrema. Por cuestiones de notación aquí se definirá la pobreza como la unión de la pobreza extrema y la pobreza moderada.

A.1.2 Agregación

Una vez que se calculan las variables dicotómicas asociadas a los cuadrantes de pobreza para cada individuo, se requiere especificar las medidas de agregación a fin de generar indicadores que midan la pobreza. En este marco de referencia, el CONEVAL define tres tipos de medidas de pobreza: de incidencia (proporción de la población), de profundidad y de intensidad (CONEVAL, 2010).

Las medidas de incidencia son las proporciones de la población pertenecientes a cada uno de los cinco cuadrantes definidos anteriormente en cada municipio, es decir,

$$p_d^{(k)} = \frac{\sum_{j=1}^{N_d} C_{dj}^{(k)}}{N_d} \times 100, \quad d = 1, \dots, D; k = 1, \dots, 6.$$

A.7

Nótese que el superíndice k se refiere a cada uno de los cuadrantes definidos en las ecuaciones A.1-A.6.

Las medidas de profundidad corresponden a la proporción promedio de carencias sociales de la población en los distintos cuadrantes de pobreza,³¹ esto es,

$$Prof_d^{(k)} = \frac{\sum_{j=1}^{N_d} C_{dj}^{(k)} \times IPS_{dj}}{\sum_{j=1}^{N_d} C_{dj}^{(k)}}, \quad d = 1, \dots, D; k = 1, 2, 3, 6.$$

A.8

En este caso, el superíndice k se refiere a los cuadrantes definidos en las ecuaciones A.1-A.3 y A.6.

Las medidas de intensidad se definen para las poblaciones siguientes: población en situación de pobreza, pobreza extrema, pobreza moderada y vulnerable por carencias.³² Sea q el número de carencias sociales, es decir, $q = 6$. Entonces,

$$Int_d^{(k)} = \frac{\left(\sum_{j=1}^{N_d} C_{dj}^{(k)} \times IPS_{dj}\right)/q}{N_d}, \quad d = 1, \dots, D; k = 1, 2, 3, 6.$$

A.9

³¹ Por cuestiones de notación se ha puesto la fórmula general de la profundidad para todos los cuadrantes de pobreza. Sin embargo, nótese que para la población no pobre y no vulnerable, así como para la población vulnerable por ingresos el número de carencias es cero. Cabe notar que la ecuación A.8 también sirve para calcular la profundidad de todas las carencias sociales.

³² La intensidad también se puede calcular para la población vulnerable por carencias sociales mediante la ecuación A.9. Nótese que si bien el subíndice se define para todos los cuadrantes, la intensidad para la población vulnerable por ingresos es cero.

Nuevamente, el superíndice k se refiere a los cuadrantes definidos en las ecuaciones A.1-A.3 y A.6. Por lo tanto, los parámetros de interés son las medidas definidas en las ecuaciones A.7-A.9.

Por otra parte, para cumplir los ordenamientos de la LGDS, el CONEVAL reporta tres indicadores asociados al contexto territorial.

- El índice de Gini se calcula mediante la técnica descrita en Deaton (1997). Para esto se ordenan a los individuos en cada municipio de acuerdo con su ingreso. Sea o_{dj} la posición del individuo j con respecto al ingreso en el municipio d de manera que la persona con mayor ingreso en el municipio d recibe la primera posición, y la persona con el menor ingreso recibe la N_d -ésima posición. El índice de Gini en el municipio d se define entonces como

$$\gamma_d = \frac{N_d + 1}{N_d - 1} - \frac{2}{N_d(N_d - 1)\bar{y}_d} \sum_{j=1}^{N_d} y_{dj}o_{dj}, \quad d = 1, \dots, D,$$

A.10

donde

$$\bar{y}_d = \frac{\sum_{j=1}^{N_d} y_{dj}}{N_d}.$$

- Por otro lado, el indicador de polarización social del municipio d se calcula con la información del índice de marginación por localidad calculado por el Consejo Nacional de Población (CONAPO). Para esto, se tiene que la población d de tamaño N_d está particionada en L_d subpoblaciones o localidades de tamaño N_{dl} , $l = 1, \dots, L_d$.

Denotemos al índice de marginación como IM , esto es, IM_{dl} es el índice de marginación de la localidad l perteneciente al municipio d . Sean P_{Ad} , P_{Md} y P_{Bd} los porcentajes de la población en el municipio d que pertenecen a localidades con índices de marginación: alto y muy alto; medio; y bajo y muy bajo, respectivamente. Es decir,

$$P_{Ad} = \frac{\sum_{l=1}^{L_d} N_{dl} \times (I(IM_{dl} = \text{"Alto"}) + I(IM_{dl} = \text{"Muy alto"}))}{N_d} \times 100,$$

$$P_{Md} = \frac{\sum_{l=1}^{L_d} N_{dl} \times (I(IM_{dl} = \text{"Medio"}))}{N_d} \times 100,$$

$$P_{Bd} = \frac{\sum_{l=1}^{L_d} N_{dl} \times (I(IM_{dl} = \text{"Bajo"}) + I(IM_{dl} = \text{"Muy bajo"}))}{N_d} \times 100.$$

Entonces, se dice que el municipio d es

- polarizado si $I(P_{Ad} > 30 \text{ y } P_{Bd} > 30 \text{ y } P_{Md} < 20) = 1$,
- no polarizado, pero con polo izquierdo si $I(P_{Ad} > 70) = 1$,
- no polarizado, pero con polo derecho si $I(P_{Bd} > 70) = 1$, o
- sin polo si

$$I\{[I(P_{Ad} > 30 \text{ y } P_{Bd} > 30 \text{ y } P_{Md} < 20) = 0] \text{ y } [I(P_{Ad} > 70) = 0] \text{ y } [I(P_{Bd} > 70) = 0]\} = 1.$$

- Finalmente, se calcula la razón entre el promedio de ingresos de la población con pobreza extrema y el promedio de ingresos de la población no pobre y no vulnerable de cada municipio, esto es,

$$R_d = \frac{\frac{\sum_{j=1}^{N_d} C_{dj}^{(2)} y_{dj}}{\sum_{j=1}^{N_d} C_{dj}^{(2)}}}{\frac{\sum_{j=1}^{N_d} C_{dj}^{(5)} y_{dj}}{\sum_{j=1}^{N_d} C_{dj}^{(5)}}}, \quad d = 1, \dots, D.$$

A.11

Nótese que todos los parámetros de interés se pueden calcular sólo si se conocen tanto el ingreso como el índice de privación social para cada individuo j del área d . Sin embargo, en la realidad no se conocen estas dos variables para todos los individuos y, por lo tanto se deben estimar. En las dos secciones siguientes se describe la metodología para estimar el ingreso para cada persona en la muestra del censo y para estimar la totalidad de parámetros de interés. En ambos casos, se asume que el Índice de Privación Social es conocido. Sin embargo, es importante recordar que en realidad éste fue estimado, como se describe en las secciones cuatro y cinco de este anexo.

A.2 Método ELL para la estimación del ingreso

En esta sección se describe brevemente el método ELL usado para la estimación del ingreso con la información disponible para la medición de la pobreza municipal en México en 2010. Para más detalles teóricos sobre este modelo ver Elbers y colaboradores (2002, 2003) y Demombynes y colaboradores (2008).

Como ya se ha mencionado, y sin pérdida de generalidad, se omitirá la referencia a los estratos de municipios, pues la metodología se aplica de forma independiente en cada estrato. También, se hace notar que el índice de privación social (IPS) será considerado como una variable dada.

La idea general es la siguiente: los datos del MCS-ENIGH se ajustan a un modelo en donde el logaritmo del ingreso es la variable dependiente. Para esto, se restringe el conjunto de las variables independientes, de manera que se encuentren tanto en el MCS-ENIGH como en la

muestra del censo. Este modelo será entonces usado para estimar el ICTPC para cada observación en la muestra del censo.

La selección de variables se restringe a aquellas variables que tienen la misma distribución tanto en el MCS-ENIGH como en la muestra del censo (Elbers y colaboradores, 2003). Otras variables a nivel municipal obtenidas tanto de la muestra del censo como de otras fuentes de información también son usadas como variables independientes. Sea $\mathbf{X} = (X_1, X_2, \dots, X_{p-1}, X_p)$ el conjunto de las variables seleccionadas para ajustar los modelos.

Sea y_{dh} la medición del ICTPC del hogar h en el municipio d en la muestra MCS-ENIGH. La estimación del ICTPC se realiza a nivel hogar; posteriormente, este ingreso es utilizado para calcular el ICTPC de los individuos pertenecientes al hogar.

Consideremos que hay un total de N^* hogares en la muestra del censo y, como antes, que la muestra está particionada en D áreas pequeñas o municipios (en este caso, $D = 2,456$ municipios). Cada área pequeña tiene tamaño N_1^*, \dots, N_D^* . Consideremos también que el total de individuos contenidos en los N^* hogares de la muestra del censo son N^{**} y que dentro de cada municipio son $N_1^{**}, \dots, N_D^{**}$. De manera similar, supóngase que hay un total de n^* hogares en el MCS-ENIGH y que éste se divide en D_s áreas pequeñas o municipios; en el MCS-ENIGH no se tiene información de todos los municipios. Por lo tanto, $D_s \leq D$. Cada municipio en el MCS-ENIGH tiene tamaño de muestra $n_1^*, \dots, n_{D_s}^*$.

Se ajusta un modelo lineal con interceptos aleatorios a nivel municipio para estimar el logaritmo del ICTPC de los hogares, es decir,

$$\ln(y_{dh}) = E(\ln(y_{dh})|\mathbf{x}_{dh}^t) + u_{dh} = \mathbf{x}_{dh}^t \boldsymbol{\beta} + u_{dh}, \quad d = 1, \dots, D_s; h = 1, \dots, n_d^*$$

A.12

donde $\mathbf{x}_{dh} = (x_{1dh}, \dots, x_{pdh})$ es el vector con los valores observados en las p variables independientes en el hogar h del municipio d en el MCS-ENIGH. El vector \mathbf{u} tiene distribución $\mathcal{F}(\mathbf{0}, \boldsymbol{\Sigma})$,³³ donde

$$u_{dh} = \eta_d + \varepsilon_{dh},$$

$$\eta_d \sim \mathcal{F}(0, \sigma_\eta^2), \quad \varepsilon_{dh} \sim \mathcal{F}(0, \sigma_{\varepsilon_{dh}}^2),$$

$$\text{Cov}(\eta_d, \eta_{d'}) = 0 \quad \forall d \neq d',$$

$$\text{Cov}(\eta_d, \varepsilon_{d'h}) = 0 \quad \forall d, d', h,$$

$$\text{Cov}(\varepsilon_{dh}, \varepsilon_{d'h'}) = 0 \quad \forall d \neq d' \text{ o } h \neq h'.$$

³³ $\mathcal{F}(\mathbf{0}, \boldsymbol{\Sigma})$ denota una distribución cualquiera con media $\mathbf{0}$ y matriz de varianzas-covarianzas $\boldsymbol{\Sigma}$.

η y ε son dos variables independientes de las variables consideradas en el vector x_{dh} . Nótese que los η_d son igualmente distribuidos (misma varianza), no así los ε_{dh} . Con estos supuestos Σ , una matriz de tamaño $n^* \times n^*$, se expresa como sigue

$$\Sigma = \begin{pmatrix} \sigma_\eta^2 + \sigma_{\varepsilon_{11}}^2 & \cdots & \sigma_\eta^2 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \cdots & \vdots & \ddots & \vdots \\ \sigma_\eta^2 & \cdots & \sigma_\eta^2 + \sigma_{\varepsilon_{1n_1}^*} & \cdots & 0 & \cdots & 0 \\ \vdots & & \vdots & \ddots & \vdots & & \vdots \\ 0 & \cdots & 0 & \cdots & \sigma_\eta^2 + \sigma_{\varepsilon_{D_s1}}^2 & \cdots & \sigma_\eta^2 \\ \vdots & \ddots & \vdots & \cdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & \sigma_\eta^2 & \cdots & \sigma_\eta^2 + \sigma_{\varepsilon_{D_s n_{D_s}^*}}^2 \end{pmatrix}.$$

A.13

El proceso de estimación es el siguiente.

- 1- Se obtiene la estimación $\hat{\beta}$ de β por el método de mínimos cuadrados ordinarios.
- 2- Sean \hat{u}_{dh} los errores de la regresión estimada en el paso anterior, es decir, $\hat{u}_{dh} = \ln(y_{dh}) - x_{dh}^t \hat{\beta}$.
- 3- Se obtienen residuales e_{dh} a partir de la descomposición siguiente:

$$\hat{u}_{dh} = \hat{u}_d + (\hat{u}_{dh} - \hat{u}_d) = \hat{\eta}_d + e_{dh},$$

donde

$$\hat{u}_d = \frac{\sum_{h=1}^{n_d^*} \hat{u}_{dh}}{n_d^*}.$$

- 4- Con la finalidad de modelar la heterogeneidad de la varianza, el modelo permite ajustar un modelo lineal a la transformación logit de los e_{dh}^2 como sigue

$$\ln \left[\frac{e_{dh}^2}{A - e_{dh}^2} \right] = \mathbf{z}_{dh}^T \boldsymbol{\alpha} + r_{dh},$$

donde $A = (1.05) \cdot \max\{e_{dh}^2\}$ y las variables en el vector \mathbf{z} son las mismas variables \mathbf{x} o una función de éstas. De aquí se obtienen los estimadores $\hat{\boldsymbol{\alpha}}$ y $\hat{\Sigma}_\alpha$.

- 5- Se calculan estimaciones para $\sigma_{\varepsilon_{dh}}^2$ como sigue,

sea $B = e^{\mathbf{z}_{dh}^T \hat{\boldsymbol{\alpha}}}$, entonces

$$\hat{\sigma}_{\varepsilon_{dh}}^2 \approx \left[\frac{AB}{1+B} \right] + \frac{1}{2} \hat{\sigma}_r^2 \left[\frac{AB(1-B)}{(1+B)^3} \right].$$

A.14

Los parámetros A y B representan las cotas para estimar la heterogeneidad de la varianza de E_{dh} .

- 6- Se debe estimar σ_{η}^2 , por ejemplo, de manera no paramétrica usando todos los valores observados $\hat{\eta}_d$ o bien usando la siguiente expresión del anexo 1 de Elbers y colaboradores (2002). Sean:

$$\hat{u}_d = \frac{\sum_{h=1}^{n_d^*} \hat{u}_{dh}}{n_d^*},$$

$$\hat{u}_{..} = \frac{\sum_{d=1}^{D_s} \sum_{h=1}^{n_d^*} \hat{u}_{dh}}{n^*},$$

$$e_d = \frac{\sum_{h=1}^{n_d^*} e_{dh}}{n_d^*},$$

$$\hat{t}_d^2 = \frac{1}{n_d^*(n_d^* - 1)} \sum_{h=1}^{n_d^*} (e_{dh} - e_d)^2.$$

Entonces,

$$\hat{\sigma}_{\eta}^2 = \max \left(\frac{\sum_{d=1}^{D_s} w_d (\hat{u}_d - \hat{u}_{..})^2}{\sum_{d=1}^{D_s} w_d (1 - w_d)} - \frac{\sum_{d=1}^{D_s} w_d (1 - w_d) \hat{t}_d^2}{\sum_{d=1}^{D_s} w_d (1 - w_d)}, \quad 0 \right)$$

A.15

donde $w_d, d = 1, \dots, D$, son números positivos tales que $\sum_{d=1}^{D_s} w_d = 1$.

- 7- Se estima la matriz Σ descrita en la ecuación A.13, sustituyendo $\sigma_{\varepsilon_{dh}}^2$ por $\hat{\sigma}_{\varepsilon_{dh}}^2$ y σ_{η}^2 por $\hat{\sigma}_{\eta}^2$, que son calculadas con base en las expresiones de las ecuaciones A.14 y A.15.
- 8- Se estima el modelo de la ecuación A.12 por mínimos cuadrados generalizados y usando la matriz $\hat{\Sigma}$ estimada en 7. De aquí se obtienen $\hat{\beta}$ y $\hat{\Sigma}_{\beta}$.

En este procedimiento de estimación se obtienen los siguientes estimadores:

$$\hat{\beta}, \hat{\Sigma}_{\beta}, \hat{\alpha}, \hat{\Sigma}_{\alpha}, \hat{\sigma}_{\eta}^2, \{\hat{\sigma}_{\varepsilon_{dh}}^2\}_{d,h},$$

que serán usados en la siguiente sección para la estimación de los parámetros de interés que miden multidimensionalmente la pobreza a nivel municipal.

A.3 Método ELLG para la estimación de las medidas de incidencia, profundidad e intensidad de la pobreza, así como los indicadores de cohesión social

En esta sección se describe el procedimiento de remuestreo (*bootstrap*) del método ELLG utilizado para la estimación de las medidas de incidencia, profundidad e intensidad de la pobreza descritas en las ecuaciones A.7 a A.9 (que se ha denotado ELLG). Se describen también los estimadores para la estimación de los indicadores de cohesión social descritos en las ecuaciones A.10 y A.11. Todas estas medidas dependen del ingreso y, en algunos casos, del Índice de Privación Social (*IPS*).

Consideremos que para todos y cada uno de los individuos de la muestra del censo se ha observado el *IPS*. Sea X la matriz de diseño donde cada columna corresponde a las variables independientes usadas en el ajuste del modelo en la ecuación A.12 y los renglones corresponden a los hogares. Sea f_{dj} el factor de expansión en la muestra del censo del individuo j del municipio d .

Sea L el número total de simulaciones de remuestreo ($l = 1, \dots, L$). Entonces, el procedimiento de estimación puntual y de los errores cuadráticos medios es el siguiente:

En cada estrato se realiza lo siguiente.

- 1- Sea $l = 1$ la primera simulación o réplica.
- 2- Se simula el ingreso para cada hogar de la muestra del censo de la siguiente manera:
 - a) Se genera un vector aleatorio de una distribución $N(\hat{\beta}, \hat{\Sigma}_{\beta})$; sea éste $\tilde{\beta}^{(l)}$;
 - b) Se genera un vector aleatorio de una distribución $N(\hat{\alpha}, \hat{\Sigma}_{\alpha})$; sea éste $\tilde{\alpha}^{(l)}$;
 - c) Se generan D números aleatorios independientes de la distribución $\mathcal{F}(0, \hat{\sigma}_{\eta}^2)$, donde \mathcal{F} puede o no ser una distribución normal. Sean estos $\tilde{\eta}_1^{(l)}, \dots, \tilde{\eta}_D^{(l)}$;
 - d) Se generan números aleatorios independientes de las distribuciones $\mathcal{F}(0, \hat{\sigma}_{\varepsilon_{11}}^2)^{(l)}, \dots, \mathcal{F}(0, \hat{\sigma}_{\varepsilon_{1N_1^*}}^2)^{(l)}, \dots, \mathcal{F}(0, \hat{\sigma}_{\varepsilon_{D1}}^2)^{(l)}, \dots, \mathcal{F}(0, \hat{\sigma}_{\varepsilon_{DN_1^*}}^2)^{(l)}$,

donde \mathcal{F} puede o no ser una distribución normal y donde $\hat{\sigma}_{\varepsilon_{dh}}^2$ se calcula evaluando la ecuación A.14 y considerando que $B = e^{z_{dh}^T \tilde{\alpha}^{(l)}}$.

Sean, $\tilde{\varepsilon}_{11}^{(l)}, \dots, \tilde{\varepsilon}_{1N_1^*}^{(l)}, \dots, \tilde{\varepsilon}_{D1}^{(l)}, \dots, \tilde{\varepsilon}_{DN_1^*}^{(l)}$, respectivamente.

- e) Se evalúa la función siguiente para obtener un valor simulado del ingreso en cada hogar

$$\ln(y_{dh}^{(l)}) = \mathbf{x}_{dh}^t \tilde{\boldsymbol{\beta}}^{(l)} + \tilde{\eta}_1^{(l)} + \tilde{\varepsilon}_{dh}^{(l)}, \quad d = 1, \dots, D; h = 1, \dots, N_d^*.$$

Esto es, se obtienen los siguientes valores del ingreso para cada hogar en los municipios:

$$y_{11}^{(l)}, \dots, y_{1N_1^*}^{(l)}, \dots, y_{D1}^{(l)}, \dots, y_{DN_D^*}^{(l)}.$$

- 3- A cada individuo en la muestra del censo se le imputa el ICTPC simulado $y^{(l)}$ correspondiente a su hogar. De manera que para cada persona se obtiene un valor simulado del ingreso, esto es, se estiman los valores $y_{11}^{(l)}, \dots, y_{1N_1^*}^{(l)}, \dots, y_{D1}^{(l)}, \dots, y_{DN_D^*}^{(l)}$.
- 4- Para cada individuo en la muestra del censo se calculan las variables dicotómicas: $C_{dj}^{(1)(l)}$, $C_{dj}^{(2)(l)}$, $C_{dj}^{(3)(l)}$, $C_{dj}^{(4)(l)}$, $C_{dj}^{(5)(l)}$, $C_{dj}^{(6)(l)}$, sustituyendo y_{dj} por $y_{dj}^{(l)}$ en las ecuaciones A.1 a A.6, respectivamente.
- 5- Si $l < L$, entonces $l = l + 1$ y se repite el proceso desde el paso 2, generando cada vez números aleatorios independientes a las simulaciones anteriores. En caso de que $l = L$, se pasa al paso 6.
- 6- Una vez realizado el proceso anterior para todos los estratos, se realiza el proceso de calibración con la metodología explicada en la sección A.6. La calibración se realiza para cada entidad federativa por separado. Las variables de la muestra del censo usadas para la calibración son las siguientes:
 - a) Indicador de carencia por rezago educativo, I_{rezedu} ,
 - b) Indicador de carencia por acceso a los servicios de salud, I_{asalud} ,
 - c) Indicador de carencia por acceso a la seguridad social, I_{segsoc} ,
 - d) Indicador de carencia por calidad y espacios de la vivienda, I_{cv} ,
 - e) Indicador de carencia por acceso a los servicios básicos en la vivienda, I_{sbv} ,
 - f) Indicador de carencia por acceso a la alimentación, I_{ali} .
 - g) Promedio de las simulaciones del indicador de ingreso inferior a la línea de bienestar mínimo, P_{lbmin} , calculada para cada individuo a partir de las L replicas como sigue.

$$P_{lb_{min_{dj}}} = \frac{\sum_{l=1}^L P_{lb_{min_{dj}}}^{(l)}}{L},$$

donde

$$P_{lb_{min_{dj}}}^{(l)} = I(y_{dj}^{(l)} < lb_{min_{dj}}), l = 1, \dots, L.$$

- h) Promedio de las simulaciones del indicador de ingreso inferior a la línea de bienestar, P_{lb} , calculada para cada individuo a partir de las L replicas como sigue:

$$P_{lb_{dj}} = \frac{\sum_{l=1}^L P_{lb_{dj}}^{(l)}}{L},$$

donde

$$P_{lb_{dj}}^{(l)} = I(y_{dj}^{(l)} < lb_{dj}), l = 1, \dots, L.$$

- i) Indicador de al menos una carencia, $I_{\geq 1car}$, calculada para cada individuo como sigue:

$$I_{\geq 1car_{dj}} = I(IPS_{dj} > 0).$$

- j) Indicador de al tres o más carencias, $I_{\geq 3car}$, calculada para cada individuo como sigue:

$$I_{\geq 3car_{dj}} = I(IPS_{dj} \geq 3).$$

- k) Promedio de las simulaciones del indicador de situación de pobreza extrema, $C^{(2*)}$, calculada para cada individuo a partir de las L replicas como sigue:

$$C_{dj}^{(2*)} = \frac{\sum_{l=1}^L C_{dj}^{(2*)^{(l)}}}{L}.$$

A.16

- l) Promedio de las simulaciones del indicador de situación de pobreza, $C^{(6*)}$, calculada para cada individuo a partir de las L replicas como sigue:

$$C_{dj}^{(6*)} = \frac{\sum_{l=1}^L C_{dj}^{(6*)^{(l)}}}{L}.$$

- m) Una variable de valor constante uno para cada individuo que sirve para poder representar a la población total.

Estas variables se utilizan para calibrar a los totales de los indicadores correspondientes a cada una de las dimensiones comentadas en los incisos a nivel estatal, que son calculados

con el MCS-ENIGH.³⁴ En la sección A.6, de calibración, se comentará con mayor detalle este proceso.

Dada la composición de las cien simulaciones de pobreza y pobreza extrema para cada persona, no era posible calibrar para cien poblaciones de pobreza (una por cada simulación), cien poblaciones de pobreza extrema, mas los otros indicadores a calibrar. Por ello, se obtuvo el promedio de las indicadores (incisos g, h) para tener un indicador de pobreza y un indicador de pobreza extrema para cada individuo que mantuviera los resultados de los indicadores por municipio con los mismos resultados (en la ecuación A.19 se demuestra la equivalencia).

Los factores de expansión originales de la muestra del censo que sirven de insumo se denotan como f_{dj} para el individuo j del municipio d . Finalmente, los factores calibrados (sección A.6) se denotan como f_{dj}^{cal} para el individuo j del municipio d .

7- Para cada réplica simulada de la muestra del censo l , $l = 1, \dots, L$, se calculan las siguientes medidas agregadas de pobreza, según las fórmulas siguientes:

- Incidencia

$$p_d^{(k)(l)} = \frac{\sum_{j=1}^{N_d^{**}} f_{dj}^{cal} C_{dj}^{(k)(l)}}{\sum_{j=1}^{N_d^{**}} f_{dj}^{cal}} \times 100, \quad d = 1, \dots, D; k = 1, \dots, 6.$$

A.17

- Profundidad

$$Prof_d^{(k)(l)} = \frac{\sum_{j=1}^{N_d^{**}} f_{dj}^{cal} C_{dj}^{(k)(l)} \times IPS_{dj}}{\sum_{j=1}^{N_d^{**}} f_{dj}^{cal} C_{dj}^{(k)(l)}}, \quad d = 1, \dots, D; k = 1, 2, 3, 6.$$

- Intensidad

$$Int_d^{(k)(l)} = \frac{\left(\sum_{j=1}^{N_d^{**}} f_{dj}^{cal} C_{dj}^{(k)(l)} \times IPS_{dj} \right) / q}{\sum_{j=1}^{N_d^{**}} f_{dj}^{cal}}, \quad d = 1, \dots, D; k = 1, 2, 3, 6.$$

8- Para cada réplica simulada de la muestra del censo l , $l = 1, \dots, L$, se calculan indicadores de contexto territorial como sigue:

- Índice de Gini

³⁴ Como se comentó en la sección 6.2 se calibra a la información del MCS-ENIGH por tres razones principales: (1) por ser el instrumento que se diseñó para medir la pobreza; (2) el CONEVAL calculó y dio a conocer la pobreza por entidad federativa previamente a la medición de la pobreza por municipios y debía ser congruente, y (3) dado que la pobreza por entidad federativa se calcula cada dos años, para ser consistentes con la fuente de la serie bianual.

Sea $o_{dj}^{(l)}$ el orden con respecto al ingreso en el municipio d de la persona j en la muestra del censo, de manera que la persona que tiene el mayor ingreso en la muestra del censo en el municipio d toma el valor 1 y el de menor el valor N_d^{**} .

Se construye un orden que considera el factor de expansión, pues cada individuo j del municipio d en la muestra del censo representa a f_{dj}^{cal} individuos. De manera que la primera persona representada por el individuo j del municipio d de la muestra del censo tendrá el orden $o_{dj}^{*(l)}$ siguiente:

$$o_{dj}^{*(l)} = 1 \text{ si } o_{dj}^{(l)} = 1,$$

$$o_{dj}^{*(l)} = \sum_{k=1}^{N_d^{**}} I(o_{dk}^{(l)} < o_{dj}^{(l)}) f_{dk}^{cal} + 1 \text{ si } o_{dj}^{(l)} > 1.$$

El orden medio de los individuos que representa el individuo j del municipio d es

$$\frac{o_{dj}^{*(l)} + (o_{dj}^{*(l)} + f_{dj}^{cal} - 1)}{2} = o_{dj}^{*(l)} + \frac{1}{2}(f_{dj}^{cal} - 1).$$

De manera que el índice de Gini se calcula como

$$\gamma_d^{(l)} = \frac{\hat{N}_d + 1}{\hat{N}_d - 1} - \frac{2}{\hat{N}_d(\hat{N}_d - 1)\bar{y}_d^{(l)}} \sum_{j=1}^{N_d^{**}} f_{dj}^{cal} y_{dj}^{(l)} \left(o_{dj}^{*(l)} + \frac{1}{2}(f_{dj}^{cal} - 1) \right), \quad d = 1, \dots, D,$$

donde

$$\hat{N}_d = \sum_{j=1}^{N_d^{**}} f_{dj}^{cal} \text{ y } \bar{y}_d^{(l)} = \frac{\sum_{j=1}^{N_d^{**}} f_{dj}^{cal} y_{dj}^{(l)}}{\hat{N}_d}.$$

- La razón entre el promedio de ingresos en la población con pobreza extrema y el promedio de ingresos en la población no pobre y no vulnerable de cada municipio está dada por:

$$R_d^{(l)} = \frac{\sum_{j=1}^{N_d^{**}} f_{dj}^{cal} C_{dj}^{(2)(l)} y_{dj}^{(l)}}{\sum_{j=1}^{N_d^{**}} f_{dj}^{cal} C_{dj}^{(2)(l)}} \cdot \frac{\sum_{j=1}^{N_d^{**}} f_{dj}^{cal} C_{dj}^{(5)(l)} y_{dj}^{(l)}}{\sum_{j=1}^{N_d^{**}} f_{dj}^{cal} C_{dj}^{(5)(l)}}, \quad d = 1, \dots, D.$$

- 9- Finalmente, las estimaciones puntuales de las medidas agregadas de incidencia, profundidad e intensidad de la pobreza y de los errores cuadráticos medios (ECM) se calculan de la manera siguiente:

a) Medidas de pobreza:

- Incidencia

$$\hat{p}_d^{(k)} = \frac{\sum_{l=1}^L p_d^{(k)(l)}}{L},$$

A.18

$$\widehat{ECM}(\hat{p}_d^{(k)}) = \frac{\sum_{l=1}^L (p_d^{(k)(l)} - \hat{p}_d^{(k)})^2}{L}, \quad d = 1, \dots, D; k = 1, \dots, 6.$$

Cabe mencionar que esta es la formulación general para calcular cualquiera de los indicadores de incidencia para cada cuadrante de pobreza. Sin embargo, únicamente fue necesario estimar la proporción de la población en situación de pobreza y pobreza extrema. Los otros cuadrantes, como se mencionó en la sección 6 del presente documento, se calculan por complemento.

- Profundidad

$$\widehat{Prof}_d^{(k)} = \frac{\sum_{l=1}^L Prof_d^{(k)(l)}}{L},$$

$$\widehat{ECM}(\widehat{Prof}_d^{(k)}) = \frac{\sum_{l=1}^L (Prof_d^{(k)(l)} - \widehat{Prof}_d^{(k)})^2}{L}, \quad d = 1, \dots, D; k = 1, 2, 3, 6.$$

- Intensidad

$$\widehat{Int}_d^{(k)} = \frac{\sum_{l=1}^L Int_d^{(k)(l)}}{L},$$

$$\widehat{ECM}(\widehat{Int}_d^{(k)}) = \frac{\sum_{l=1}^L (Int_d^{(k)(l)} - \widehat{Int}_d^{(k)})^2}{L}, \quad d = 1, \dots, D; k = 1, 2, 3, 6.$$

b) Indicadores de contexto territorial:

- Índice de Gini

$$\hat{\gamma}_d = \frac{\sum_{l=1}^L \gamma_d^{(l)}}{L},$$

$$\widehat{ECM}(\hat{\gamma}_d) = \frac{\sum_{l=1}^L (\gamma_d^{(l)} - \hat{\gamma}_d)^2}{L}, \quad d = 1, \dots, D.$$

- La razón entre el promedio de ingresos en la población con pobreza extrema y el promedio de ingresos en la población no pobre y no vulnerable de cada municipio.

$$\hat{R}_d = \frac{\sum_{l=1}^L R_d^{(l)}}{L},$$

$$\widehat{ECM}(\hat{R}_d) = \frac{\sum_{l=1}^L (R_d^{(l)} - \hat{R}_d)^2}{L}, \quad d = 1, \dots, D.$$

Mediante este desarrollo matemático, se cuenta con todos los indicadores de pobreza y de cohesión social para los municipios de México.

Finalmente, como se mencionó en 6 es necesario demostrar que es equivalente utilizar el promedio de las simulaciones del indicador de pobreza y pobreza extrema para “representar” a los individuos en el proceso de calibración. La ecuación para las estimaciones de los cuadrantes de pobreza están en A.18. Esta ecuación se puede reescribir sustituyendo A.17 como:

$$\hat{p}_d^{(k)} = \frac{\sum_{j=1}^{N_d^{**}} f_{dj}^{cal} \sum_{l=1}^L \frac{C_{dj}^{(k)(l)}}{L}}{\sum_{j=1}^{N_d^{**}} f_{dj}^{cal}} \times 100, \quad d = 1, \dots, D; k = 1, \dots, 6.$$

Por la ecuación A.16 de pobreza, se sabe que:

$$C_{dj}^{(2*)} = \frac{\sum_{l=1}^L C_{dj}^{(6)(l)}}{L}.$$

Es decir, el estimador de pobreza extrema está dado por la fórmula:

$$\hat{p}_d^{(2)} = \frac{\sum_{j=1}^{N_d^{**}} f_{dj}^{cal} C_{dj}^{(6*)}}{\sum_{j=1}^{N_d^{**}} f_{dj}^{cal}} \times 100.$$

A.19

Nótese que se pueden derivar expresiones análogas para el estimador de pobreza extrema ($k = 2$) o para cualquier otro indicador para el cual se hayan generado réplicas de simulación. Ello demuestra que es equivalente matemáticamente utilizar el promedio de las simulaciones, para obtener los estimadores de los cuadrantes de pobreza por municipios.

A.4 Modelos de regresión logística usados para la estimación de los indicadores de carencias por acceso a la seguridad social y por acceso a la alimentación

En esta sección se describen los modelos de regresión logística usados para estimar los indicadores de carencia por acceso a la seguridad social y el indicador de carencia por acceso a la alimentación. Estos indicadores no se pueden construir con la información disponible en la muestra del censo. Por lo tanto, se utilizó el MCS-ENIGH 2010 para modelar la probabilidad de que un individuo u hogar presentara la carencia por acceso a la seguridad social o bien la carencia por acceso a la alimentación.

Como se hizo para el caso del ingreso, se restringió el conjunto de variables independientes de los modelos de manera que ambas estuvieran disponibles tanto en el MCS-ENIGH como en la muestra del censo. Estas fueron las variables independientes de los modelos para estimar el indicador de carencia por acceso a la seguridad social y el indicador de carencia por acceso a la alimentación.

En la selección de las variables se verificó que las variables explicativas de los modelos tuvieran la misma distribución en el MCS-ENIGH y en la muestra del censo. Otras variables a nivel municipal obtenidas tanto de la muestra del censo como de otras fuentes de información también fueron utilizadas como variables predictoras. Sea $W = (W_1, W_2, \dots, W_{p-1}, W_p)$ el conjunto de variables independientes seleccionadas para ajustar los modelos. Cabe mencionar que un subconjunto de las variables en W son variables con información propia de los individuos u hogares y otro subconjunto son variables con información municipal.

Al igual que en las secciones anteriores, se ha omitido en la notación la referencia a los grupos de municipios o estratos. La metodología se aplica de forma independiente en cada uno de estos estratos.

Se considera que hay un total de N^* hogares en la muestra del censo y que ésta se encuentra particionada en D áreas pequeñas o municipios. Cada área pequeña tiene tamaño N_1^*, \dots, N_D^* . Se considera también que el total de individuos contenidos en los N^* hogares de la muestra del censo es N^{**} y que en cada municipio hay $N_1^{**}, \dots, N_D^{**}$ individuos. De manera similar, se ha considerado que hay un total de n^* hogares en el MCS-ENIGH y que la muestra del módulo se divide en D_s municipios. En el MCS-ENIGH no se tiene información de todos los municipios, por lo que $D_s \leq D$. La muestra de cada municipio en el MCS-ENIGH tiene tamaño $n_1^*, \dots, n_{D_s}^*$. El total de individuos en los n^* hogares del MCS-ENIGH es n^{**} . En cada municipio hay $n_1^{**}, \dots, n_{D_s}^{**}$ individuos representados.

A.4.1 Modelo de regresión logística para el indicador de carencia por acceso a la alimentación

Sea $I_{alim,dh}$ la variable dicotómica observada en el MCS-ENIGH que vale uno si el hogar h del municipio d presenta carencia por acceso a la alimentación y cero en caso contrario. Se modeló la

probabilidad (Gosh, 1998) de que un hogar presentará la carencia por acceso a la alimentación como sigue:

$$P(I_{alim_{dh}} = 1 | \mathbf{w}_{dh}) = p_{alim_{dh}} = \frac{\exp^{w_{dh}^t \beta}}{1 + \exp^{w_{dh}^t \beta}}, \quad h = 1, \dots, n_d^*; d = 1, \dots, D_s$$

Para estimar los parámetros se supuso que la variable $I_{alim_{dh}}$ sigue una distribución Bernoulli con parámetro $p_{alim_{dh}}$. Esto es, se calculó la función de verosimilitud y se usaron métodos iterativos para encontrar los estimadores máximos verosímiles (*Stata* utiliza el método de Newton-Raphson modificado, el Berndt-Hall-Hall-Hausman, Davidon-Fletcher-Powell, Broyden-Fletcher-Goldfarb-Shanno).³⁵

Una vez que se tenía el estimador $\hat{\beta}$, se estimaron las probabilidades para cada hogar en la muestra del censo evaluando la función:

$$\hat{p}_{alim_{dh}} = \frac{\exp^{w_{dh}^t \hat{\beta}}}{1 + \exp^{w_{dh}^t \hat{\beta}}}, \quad h = 1, \dots, N_d^*; d = 1, \dots, D.$$

Posteriormente, se imputó la probabilidad $\hat{p}_{alim_{dh}}$ a todos los individuos pertenecientes al hogar h , de manera que cada individuo en la muestra del censo tenía una probabilidad $\hat{p}_{alim_{dj}}$ de ser carente, $j = 1, \dots, N_d^{**}$; $d = 1, \dots, D$. A partir de las probabilidades $\hat{p}_{alim_{dj}}$ se predijeron valores cero o uno para cada individuo dependiendo de un valor de corte c_d , $d = 1, \dots, D$, de manera que

$$\hat{I}_{alim_{dj}} = \begin{cases} 1, & \hat{p}_{alim_{dj}} > c_d \\ 0, & \hat{p}_{alim_{dj}} \leq c_d \end{cases}, \quad j = 1, \dots, N_d^{**}; d = 1, \dots, D.$$

El valor de corte se obtuvo al estimar la probabilidad de presentar carencia por acceso a la alimentación en el municipio. Se calculó como $c_d = \frac{\sum_{j=1}^{N_d^{**}} f_{dj} \hat{p}_{alim_{dj}}}{\sum_{j=1}^{N_d^{**}} f_{dj}}$, $d = 1, \dots, D$, donde f_{dj} era el factor de expansión en la muestra del censo del individuo j del municipio d .

A.4.2 Modelo de regresión logística mixta para el indicador de carencia por acceso a la seguridad social

Sea $I_{segsoc_{dj}}$ la variable dicotómica observada en el MCS-ENIGH que vale uno si el individuo j del municipio d presenta carencia por acceso a la seguridad social. Se modeló la probabilidad de que un individuo presentara carencia por acceso a la seguridad social usando un modelo de regresión logística con interceptos aleatorios como sigue:

$$P(I_{segsoc_{dj}} = 1 | \mathbf{w}_{dj}) = p_{segsoc_{dj}} = \frac{\exp^{w_{dj}^t \beta + u_d}}{1 + \exp^{w_{dj}^t \beta + u_d}}, \quad j = 1, \dots, n_d^*; d = 1, \dots, D_s,$$

³⁵ Gould y colaboradores (2010).

donde u_1, \dots, u_{D_s} son los coeficientes aleatorios asociados a cada uno de los municipios. Se supuso que éstos son variables independientes y que siguen una distribución $\mathcal{F}(0, \varphi)$.

Para estimar los parámetros fijos del modelo, β , se supuso que la variable $I_{segsoc_{dj}}|u_d$ sigue una distribución Bernoulli con parámetro $p_{segsoc_{dj}}$. Así pues, tanto para la estimación de β como de $u_d, d = 1, \dots, D_s$, se usaron los algoritmos iterativos (HLM utiliza el *penalized quasi-likelihood*).³⁶

Una vez que se había obtenido el vector $\hat{\beta}$ y $\hat{u}_d, d = 1, \dots, D_s$, se estimaron las probabilidades para cada individuo en los municipios $1, \dots, D_s$ en la muestra del censo evaluando la función siguiente:

$$\hat{p}_{segsoc_{dj}} = \frac{\exp^{w_{dj}^t \hat{\beta} + \hat{u}_d}}{1 + \exp^{w_{dj}^t \hat{\beta} + \hat{u}_d}}, \quad j = 1, \dots, N_d^{**}; d = 1, \dots, D_s.$$

En la expresión anterior se supone que de los D municipios en la muestra del censo, los correspondientes a los subíndices $1, \dots, D_s$ son los que se encuentran en el MCS-ENIGH.

Por otro lado, para los municipios que no están en la muestra, se usó la expresión siguiente:

$$\hat{p}_{segsoc_{dj}} = \frac{\exp^{w_{dj}^t \hat{\beta}}}{1 + \exp^{w_{dj}^t \hat{\beta}}}, \quad j = 1, \dots, N_d^{**}; d = D_s + 1, \dots, D.$$

Finalmente, a partir de las probabilidades $\hat{p}_{segsoc_{dj}}$ se predijeron los valores cero o uno para cada individuo dependiendo de un valor de corte $c_d, d = 1, \dots, D$, de manera que

$$\hat{I}_{segsoc_{dj}} = \begin{cases} 1, & \hat{p}_{segsoc_{dj}} > c_d \\ 0, & \hat{p}_{segsoc_{dj}} \leq c_d \end{cases}, \quad j = 1, \dots, N_d^{**}; d = 1, \dots, D.$$

El valor de corte corresponde a la estimación de la probabilidad de presentar carencia por acceso a la seguridad social en el municipio. Se calculó como $c_d = \frac{\sum_{j=1}^{N_d^{**}} f_{dj} \hat{p}_{segsoc_{dj}}}{\sum_{j=1}^{N_d^{**}} f_{dj}}, d = 1, \dots, D$, donde f_{dj} era el factor de expansión en la muestra del censo del individuo j del municipio d .

Las variables estimadas \hat{I}_{segsoc} y \hat{I}_{alim} se usaron para estimar el Índice de Privación Social.

A.5 Construcción del Índice de Privación Social (IPS)

En esta sección se describe la construcción del Índice de Privación Social. Éste índice se construye para cada persona a partir de la suma de los siguientes seis indicadores asociados a las carencias sociales:

³⁶ Raudenbush y Bryk (2002).

- Indicador de rezago educativo, I_{rezedu} ,
- Indicador de carencia por acceso a los servicios de salud, I_{asalud} ,
- Indicador de carencia por acceso a la seguridad social, I_{segsoc} ,
- Indicador de carencia por calidad y espacios de la vivienda. I_{cv} ,
- Indicador de carencia por acceso a los servicios básicos en la vivienda, I_{sbv} ,
- Indicador de carencia por acceso a la alimentación, I_{ali} .

Se define el índice de privación social para el individuo IPS para el individuo j como

$$IPS_j = I_{rezedu_j} + I_{asalud_j} + I_{segsoc_j} + I_{cv_j} + I_{sbv_j} + I_{ali_j}, \quad j = 1, \dots, N,$$

donde N es el total de individuos en la población. El IPS toma valores enteros entre 0 y 6 debido a que cada indicador es dicotómico.

En la muestra del censo no es posible calcular directamente el índice de privación social, pues dos de los indicadores, I_{segsoc} y I_{ali} , no se podían calcular con la información disponible. De aquí que el índice de privación social también se estima usando los indicadores estimados \hat{I}_{segsoc} y \hat{I}_{ali} descritos en la sección anterior, de manera que si N^{**} es el total de individuos en la muestra del censo, entonces:

$$\widehat{IPS}_j = I_{rezedu_j} + I_{asalud_j} + \hat{I}_{segsoc_j} + I_{cv_j} + I_{sbv_j} + \hat{I}_{ali_j}, \quad j = 1, \dots, N^{**}.$$

Finalmente, cabe destacar que, según la metodología descrita en las secciones dos y tres de este anexo, se ha considerado a la variable \widehat{IPS} como dada. Por lo tanto, no se consideró su variación aleatoria en la estimación de los errores cuadráticos medios de las medidas de pobreza que en esas secciones se discutieron.

A.6 Calibración de los factores de expansión de la muestra del censo

En esta sección se describe la metodología usada por el CONEVAL para realizar el proceso de calibración de la muestra del censo. Se utilizan los totales de variables calculados previamente a nivel estatal con el MCS-ENIGH.

Las variables que se usaron en el proceso de calibración se describieron en la sección A.3 de este anexo (particularmente en el paso 6). Nótese que se ha omitido el subíndice sobre los estados de la República en la descripción de esta fase de la metodología. Los totales estatales que sirven para calibrar los factores de expansión fueron obtenidos del MCS-ENIGH.

El método de calibración usado es el método *logit* descrito en Deville y Särndal (1992) y en Deville y colaboradores (1993). Este método tiene la ventaja de evitar pesos calibrados negativos y cambios bruscos con respecto a los pesos originales.

Sea $\mathbf{X} = (X_1, \dots, X_p)$ el vector de p variables usadas para el proceso de calibración. Sea $t_{X_i} = \sum_{j=1}^N x_{ij}$ el total de la variable $X_i, i = 1, \dots, p$, en la población (al referirnos a la población nos referimos a la población de la entidad federativa en turno). Estos totales se estiman con el MCS-ENIGH, es decir, tenemos un vector de totales estimados $\hat{t} = (\hat{t}_{X_1}, \dots, \hat{t}_{X_p})$. Los totales a los que se calibró fueron los de los indicadores básicos de pobreza: el rezago educativo, el acceso a los servicios de salud, la calidad y espacios de la vivienda, el acceso a los servicios básicos en la vivienda, el acceso a la alimentación, el acceso a la seguridad social, la población con ingreso inferior a la línea de bienestar y a la de bienestar mínimo, la población con al menos una carencia social, la población con tres o más carencias sociales, la población en pobreza y la población en pobreza extrema.

Con la muestra del censo también es posible estimar estos totales con el estimador Horwitz-Thompson. Cabe mencionar que en la sección A.3 se usó el subíndice d para referirnos a los municipios. Sin embargo, dado que en la descripción de la metodología de calibración no es necesario referirnos a éstos, aquí se omite este subíndice, de manera que denotamos como f_j a los factores de expansión de la muestra del censo. Los estimadores Horwitz-Thompson de los totales de la muestra del censo son

$$\tilde{t}_{X_i} = \sum_{j=1}^{N^{**}} f_j x_{ij}, \quad i = 1, \dots, p,$$

donde N^{**} es el total de individuos en la muestra del censo en la población de interés, es decir, la entidad federativa en turno.

En general, $\tilde{t}_{X_i} \neq \hat{t}_{X_i}, i = 1, \dots, p$. El objetivo, por lo tanto, es modificar los factores de expansión originales de la muestra del censo, f_j , de manera que los nuevos pesos f_j^{cal} cumplan la condición de que

$$\sum_{j=1}^{N^{**}} f_j^{cal} x_{ij} = \hat{t}_{X_i}, \quad i = 1, \dots, p.$$

A.18

Para resolver este problema se considera una métrica particular, como en Deville y Särndal (1992), y en Deville y colaboradores (1993), de manera que los nuevos pesos f_j^{cal} estén lo más cerca posible de los pesos f_j originales.

La métrica que se usa corresponde a la usada en el método de calibración logístico que a continuación se describe.

Sea $G(\cdot)$ una función de distancia con argumento $\frac{f_j^{cal}}{f_j}$. Supongamos que la función cumple con las propiedades siguientes:

- a. G es positiva y estrictamente convexa;
- b. $G(1) = G'(1) = 0$, y
- c. $G''(1) = 0$.

Sea $\sum_{j=1}^{N^{**}} f_j G\left(\frac{f_j^{cal}}{f_j}\right)$ la medida de distancia entre los pesos calibrados y los pesos originales en la muestra del censo. El problema descrito en la ecuación A.18 es un problema de minimización que se puede expresar como:

Minimizar, con respecto a f_j^{cal} ,

$$\sum_{j=1}^{N^{**}} f_j G\left(\frac{f_j^{cal}}{f_j}\right) - \lambda^t \left(\sum_{j=1}^{N^{**}} f_j^{cal} x_j - \hat{t} \right),$$

donde el vector de p entradas λ corresponde a los multiplicadores de Lagrange. Al diferenciar con respecto a f_j^{cal} e igualar a cero, se obtiene que

$$g\left(\frac{f_j^{cal}}{f_j}\right) - x_j^t \lambda = 0,$$

donde

$$g(x) = \frac{\partial G(x)}{\partial x}.$$

Al resolver para f_j^{cal} , se obtiene

$$f_j^{cal} = f_j F(x_j^t \lambda),$$

donde $F(u) = g^{-1}(u)$ denota a la función inversa de $g(\cdot)$.

Por lo tanto, para calcular los nuevos pesos basta con determinar el valor de λ . Esto se obtiene resolviendo las ecuaciones de calibración siguientes

$$\sum_{j=1}^{N^{**}} f_j F(x_j^t \lambda) x_j = \hat{t}.$$

El sistema de ecuaciones anterior se resuelve mediante métodos numéricos. Específicamente, para este propósito se implementó una función en el paquete estadístico R, que a su vez usa el procedimiento (*library*) *BB* de este mismo paquete (Varadhan y Gilbert, 2009). El algoritmo usado en este procedimiento está basado en el método *Barzilai-Borwein gradient* (Raydan, 1997).

Finalmente, el método de calibración logístico se define con los siguientes componentes:

- La métrica G es

$$G\left(\frac{f_j^{cal}}{f_j}\right) \propto \left(\frac{f_j^{cal}}{f_j} - L\right) \ln\left(\frac{\frac{f_j^{cal}}{f_j} - L}{1-L}\right) + \left(U - \frac{f_j^{cal}}{f_j}\right) \ln\left(\frac{U - \frac{f_j^{cal}}{f_j}}{U-1}\right), \text{ si } L < \frac{f_j^{cal}}{f_j} < U, \text{ y}$$

$$G\left(\frac{f_j^{cal}}{f_j}\right) = \infty \text{ en otro caso.}$$

- $F(u) = \frac{L(U-1)+U(1-L)e^{Au}}{U-1+(1-L)e^{Au}},$

donde

$$A = \frac{U - L}{(1 - L)(U - 1)}.$$

- $\frac{\partial F(u)}{\partial u} = F(u)[U - F(u)] \frac{A(1-L)e^{Au}}{U-1+(1-L)e^{Au}}.$

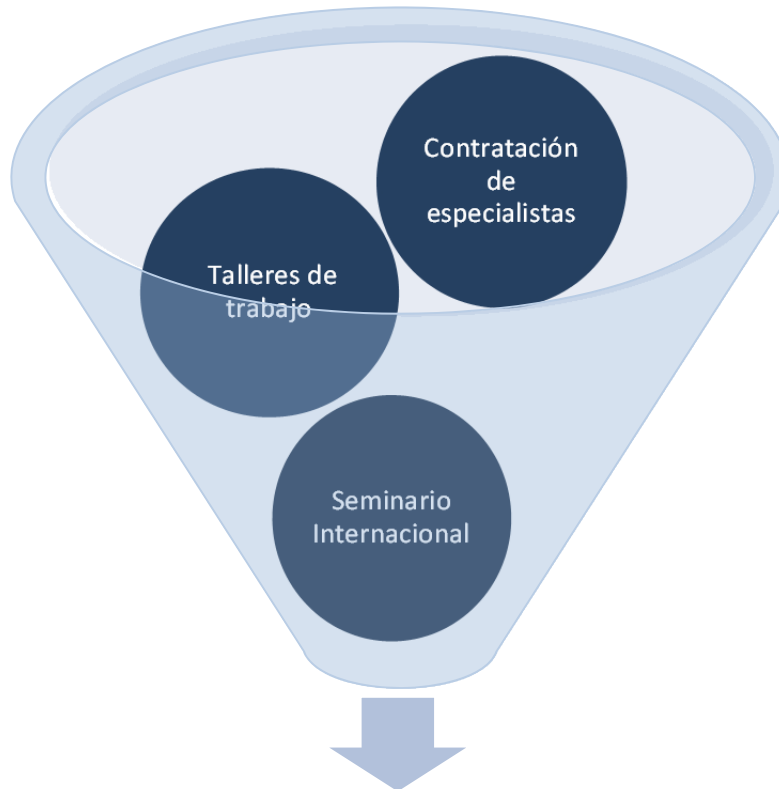
Los valores L y U en las expresiones anteriores se deben especificar *a priori* y sirven para acotar las diferencia entre los pesos calibrados y los pesos originales.

Cabe señalar que con el desarrollo matemático presentado en este anexo, se cumple con el objetivo del CONEVAL de tener las estimaciones de pobreza para todos los municipios del país como marca la LGDS, de acuerdo con lo establecido en la metodología para la medición multidimensional de la pobreza en México, y de manera consistente con las estimaciones por entidad federativa. Con ello, el CONEVAL busca contribuir con los principios de transparencia y rigor técnico al publicar el desarrollo que llevó a cabo para la medición municipal de la pobreza.

B. Mapa conceptual del proceso de investigación

En este anexo se presenta, de manera esquemática, el proceso que siguió la medición de pobreza en los municipios del país, mediante mapas conceptuales de cada una de las etapas.

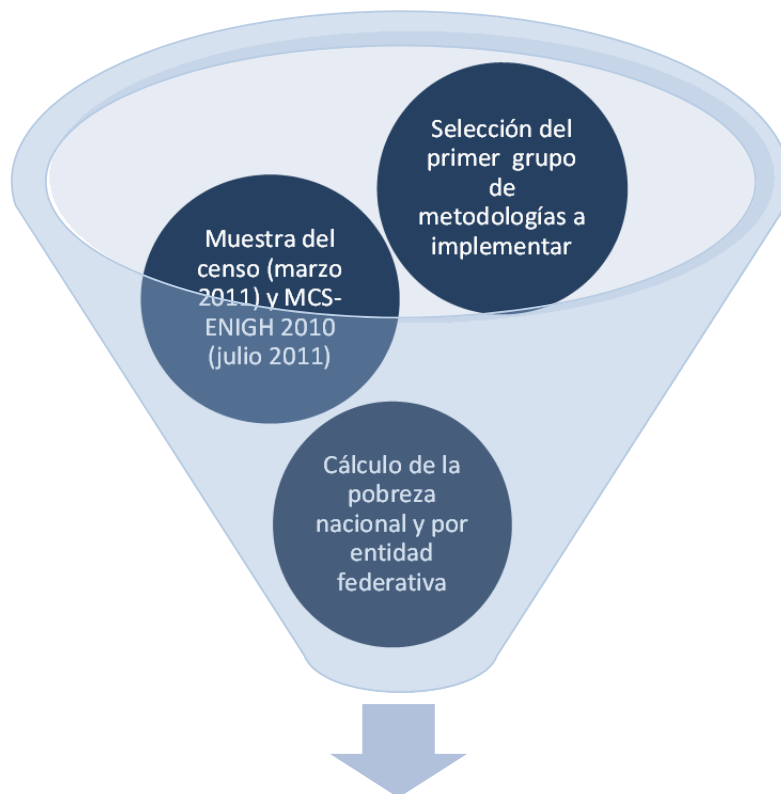
B.1 Etapa de exploración de metodologías



Metodologías desarrolladas:

- Modelo híbrido (Jae-Kwan Kim)
- Estimadores Horwitz-Thompson y análisis de conglomerados (Ignacio Méndez y Hortensia Moreno)
- Modelo bayesiano (Luis Enrique Nieto)
- Modelo mejor predictor empírico (Nikolaos Tzavidis)
- Modelo de regresión geográfica ponderada (Christopher Brunsdon)

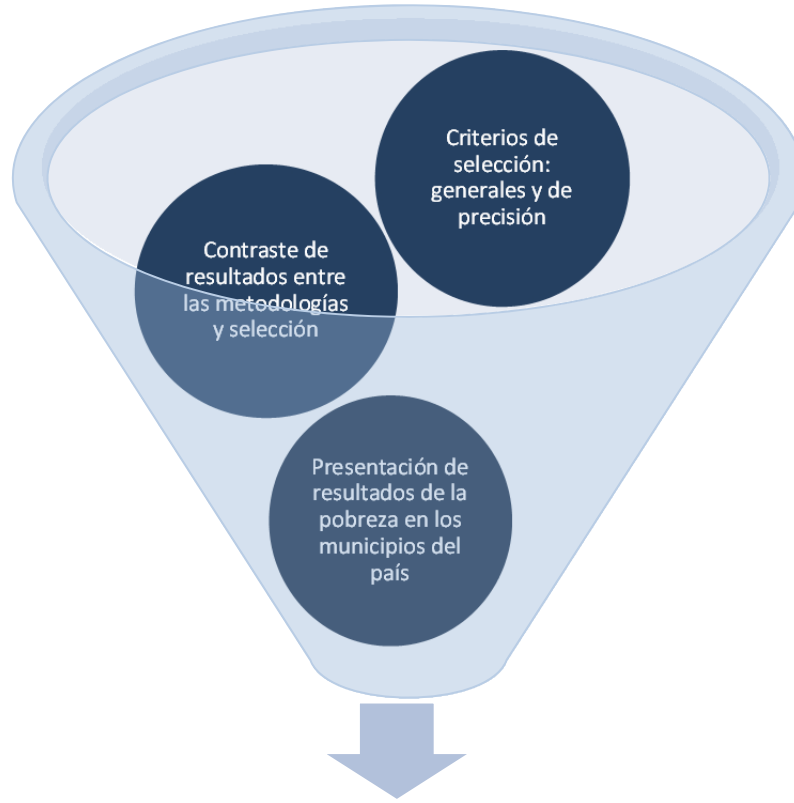
B.2 Implementación de metodologías y cálculo de la medición de pobreza nacional y por entidad federativa



Metodologías implementadas:

- Modelo bayesiano
- Modelo mejor predictor empírico
- Modelo logístico jerárquico
- Modelo multinomial
- Modelo del mejor predictor empírico adaptado y generalizado a la medición multidimensional de la pobreza
- Modelo de Elbers y colaboradores adaptado y generalizado a la medición multidimensional de la pobreza

B.3 Contraste de resultados, selección de la metodología y presentación de resultados



Resultados el 2 de diciembre de 2011

C. Información complementaria

En este anexo se presentan los cuadros con información relevante para la metodología de estimación de la pobreza municipal que están referidos en el cuerpo del documento.

Cuadro C.1
Ventajas y desventajas de las metodologías para imputación en áreas pequeñas exploradas en los años 2009-2010

Modelo	VARIABLES Y PARÁMETROS ESTIMADOS	Ventajas	Desventajas	Plataforma informática
Modelo híbrido	Totales de cada municipio para las variables: ICTPC, rezago educativo, acceso a los servicios de salud, acceso a la seguridad social, calidad y espacios de la vivienda, servicios básicos en la vivienda, acceso a la alimentación, población por debajo de la línea de bienestar y bienestar mínimo, pobreza y pobreza extrema, y los dos grupos vulnerables.	Se pueden estimar totales para variables continuas o dicotómicas. Cálculo sencillo y fácil de implementar.	No considera el uso de la información geográfica. En la construcción de los estimadores sintéticos no se considera el diseño muestral. Resultados de la aplicación no consistentes con el tamaño de la población en los municipios.	SAS
Estimadores Horvitz-Thompson de razón y mediante análisis de conglomerados	Promedio de la variable ICTPC para cada municipio, porcentaje de acceso a la alimentación para cada municipio y porcentajes de pobreza, pobreza extrema y los dos grupos vulnerables para cada municipio	Fácil de calcular. Se pueden usar rutinas ya implementadas en muchos paquetes estadísticos. Permite el cálculo considerando variables continuas o dicotómicas.	La definición de las áreas grandes que contienen a las pequeñas debe ser adecuada. Considerar el diseño muestral produce que los errores cuadráticos medios sean grandes, y por lo tanto las estimaciones son menos precisas. En el caso de proporciones, se obtienen en algunos casos intervalos fuera de rango. Las variables para obtener las estimaciones municipales se obtuvieron a nivel hogar no individuo.	STATA y JPM
Modelo bayesiano	Promedio del ICTPC para cada municipio, y proporciones de los seis indicadores de carencia a nivel individuo para cada municipio.	Estimación multivariada para variables continuas y dicotómicas. Permite el uso de la información geográfica. Es muy flexible y se puede adaptar a las 32 entidades federativas.	Intervalos de credibilidad demasiado amplios. Depende de la especificación de los parámetros iniciales y de que se sigan los supuestos distribucionales. No considera el diseño muestral.	R
Método del mejor predictor empírico	Promedio de la variable ingreso corriente total per cápita a nivel hogar para cada municipio, percentiles 25, 50 y 75 para el ingreso corriente total per cápita a nivel hogar para cada municipio, y estimaciones del índice de pobreza a nivel hogar para cada municipio.	Permite una estimación robusta de los parámetros basados en la variable continua ingreso. Además, proporciona errores cuadráticos medios pequeños. Se puede combinar con el enfoque de GWR e incluir información espacial.	No consideró el uso de información geográfica. En la construcción de los estimadores sintéticos no se considera el diseño muestral. Sólo sirve para la estimación de parámetros de variables continuas como el ingreso.	R
Modelo de regresión geográficamente ponderada	Promedio de la variable ingreso corriente total per cápita a nivel hogar para cada municipio	Permite el uso de información geográfica. Errores cuadráticos medios pequeños. Incorpora información geoespacial.	Sólo se presentó la teoría y aplicación para el caso de variables continuas y depende de supuestos distribucionales para éstas. Requiere del cálculo de distancias adecuadas, en particular, la distancia de superficie. Requiere de un costo computacional mayor a los demás métodos.	R

Fuente: elaboración del CONEVAL con base en Vargas y colaboradores (2010).

Cuadro C.2

Porcentaje de la población con rezago educativo y carencia en el acceso a los servicios de salud, según fuente de información, por entidad federativa. México, 2010

Clave	Entidad federativa	Rezago educativo					Acceso a los servicios de salud				
		MCS-ENIGH			Muestra del censo	Diferencia absoluta	MCS-ENIGH			Muestra del censo	Diferencia absoluta
		Puntual	Intervalo de confianza (95)		Puntual		Puntual	Puntual	Intervalo de confianza (95)		
			Mínimo	Máximo		Mínimo			Máximo		
01	Aguascalientes	17.2	15.8	18.5	15.1	2.1	22.0	19.9	24.1	17.1	4.8
02	Baja California	17.1	15.4	18.8	14.1	3.1	35.2	32.6	37.7	27.3	7.8
03	Baja California Sur	16.9	14.9	18.9	13.8	3.2	22.6	18.7	26.5	23.7	1.0
04	Campeche	23.9	21.3	26.6	20.8	3.1	20.8	18.6	23.0	17.6	3.2
05	Coahuila	12.1	10.8	13.4	12.9	0.8	20.1	18.0	22.1	21.1	1.0
06	Colima	18.7	17.0	20.4	18.0	0.7	17.8	16.0	19.6	16.5	1.3
07	Chiapas	35.0	32.6	37.4	32.9	2.1	36.5	32.8	40.2	43.2	6.7
08	Chihuahua	17.5	15.7	19.3	16.9	0.6	20.8	18.4	23.3	24.2	3.4
09	Distrito Federal	9.5	8.8	10.3	9.4	0.1	35.7	32.9	38.5	32.9	2.8
10	Durango	18.6	17.1	20.2	18.2	0.5	29.1	25.9	32.3	28.7	0.4
11	Guanajuato	23.6	21.8	25.4	23.5	0.1	27.1	24.2	29.9	27.3	0.2
12	Guerrero	28.3	26.2	30.3	27.9	0.4	39.6	36.2	43.0	46.1	6.6
13	Hidalgo	23.4	21.2	25.6	19.7	3.7	31.7	27.6	35.7	32.4	0.7
14	Jalisco	20.2	18.1	22.4	19.1	1.1	35.2	32.3	38.0	33.8	1.3
15	México	18.5	15.0	21.9	14.7	3.8	35.5	31.7	39.3	39.3	3.8
16	Michoacán	30.4	28.3	32.5	28.5	1.9	39.3	35.1	43.6	44.4	5.0
17	Morelos	19.4	17.8	21.1	17.3	2.1	31.7	29.3	34.2	34.1	2.4
18	Nayarit	20.1	18.8	21.5	19.7	0.4	24.1	21.0	27.3	22.3	1.9
19	Nuevo León	13.1	11.8	14.4	11.8	1.4	22.4	20.3	24.4	18.9	3.5
20	Oaxaca	30.3	27.2	33.3	29.9	0.3	39.9	35.7	44.0	44.3	4.4
21	Puebla	25.0	22.7	27.3	24.1	0.9	41.8	38.4	45.2	50.4	8.6
22	Querétaro	19.5	17.9	21.1	17.4	2.1	24.4	22.5	26.4	25.8	1.4
23	Quintana Roo	18.2	16.7	19.7	16.1	2.1	27.8	25.6	30.0	30.4	2.6
24	San Luis Potosí	22.2	20.3	24.1	20.3	1.9	21.0	18.6	23.4	24.1	3.1
25	Sinaloa	19.2	17.6	20.8	18.8	0.4	23.1	21.1	25.1	22.2	0.9
26	Sonora	14.0	12.7	15.3	14.3	0.3	26.4	23.8	29.1	24.0	2.5
27	Tabasco	19.9	18.1	21.6	18.2	1.6	25.6	22.6	28.7	23.6	2.0
28	Tamaulipas	14.5	13.0	16.0	15.9	1.4	23.1	20.8	25.5	21.9	1.3
29	Tlaxcala	15.6	14.3	16.9	15.3	0.3	35.1	32.5	37.7	36.5	1.5
30	Veracruz	26.1	24.0	28.3	26.5	0.4	36.9	33.2	40.5	40.1	3.2
31	Yucatán	24.6	23.1	26.2	23.5	1.1	22.4	20.7	24.1	22.4	0.0
32	Zacatecas	22.8	21.2	24.4	22.5	0.3	27.0	24.1	29.9	28.0	1.0
Estados Unidos Mexicanos		20.6	20.1	21.2	19.4	1.2	31.8	31.0	32.6	33.2	1.4

Nota: Estimaciones del CONEVAL con base en la muestra del Censo de Población y Vivienda 2010 y el MCS-ENIGH 2010.

Cuadro C.3
Porcentaje de la población con carencia en calidad y espacios de la vivienda y carencia de servicios básicos en la vivienda, según fuente de información, por entidad federativa. México, 2010

Clave	Entidad federativa	Calidad y espacios de la vivienda					Servicios básicos en la vivienda				
		MCS-ENIGH			Muestra del censo	Diferencia absoluta	MCS-ENIGH			Muestra del censo	Diferencia absoluta
		Puntual	Intervalo de confianza (95)		Puntual		Puntual	Puntual	Intervalo de confianza (95)		
			Mínimo	Máximo		Mínimo			Máximo		
01	Aguascalientes	6.9	3.6	10.2	6.3	0.6	3.0	1.6	4.5	4.0	1.0
02	Baja California	10.2	7.9	12.4	8.8	1.4	6.5	4.2	8.9	9.2	2.7
03	Baja California Sur	12.3	8.2	16.3	15.1	2.8	8.0	4.5	11.5	13.2	5.2
04	Campeche	22.0	18.5	25.5	24.9	2.9	20.9	16.4	25.3	22.2	1.3
05	Coahuila	4.4	2.9	5.9	6.1	1.7	5.3	3.1	7.4	7.2	1.9
06	Colima	12.1	9.8	14.4	12.9	0.9	3.2	2.0	4.3	4.1	1.0
07	Chiapas	33.2	28.2	38.3	37.1	3.8	34.1	27.6	40.5	39.8	5.8
08	Chihuahua	6.4	4.3	8.6	9.4	2.9	6.8	4.8	8.9	9.4	2.5
09	Distrito Federal	7.6	5.6	9.5	8.7	1.1	3.7	2.4	5.1	3.5	0.3
10	Durango	11.3	8.6	13.9	13.1	1.9	16.1	11.6	20.6	14.8	1.3
11	Guanajuato	9.6	7.5	11.6	11.9	2.4	13.9	10.7	17.1	17.9	4.0
12	Guerrero	40.6	37.3	43.8	41.8	1.3	40.1	34.0	46.2	50.3	10.2
13	Hidalgo	13.6	10.4	16.7	15.8	2.2	22.6	17.2	27.9	23.7	1.2
14	Jalisco	6.7	4.7	8.7	9.3	2.7	9.5	5.9	13.2	8.9	0.6
15	México	12.9	9.8	15.9	14.4	1.5	13.3	9.8	16.9	14.5	1.1
16	Michoacán	22.3	16.9	27.7	20.8	1.6	17.9	13.8	22.0	22.4	4.5
17	Morelos	15.9	13.2	18.5	16.6	0.7	14.8	10.5	19.0	17.1	2.4
18	Nayarit	12.6	9.1	16.1	13.0	0.4	10.4	6.5	14.4	16.2	5.8
19	Nuevo León	6.8	4.2	9.5	6.4	0.5	2.8	1.7	3.9	5.5	2.7
20	Oaxaca	34.1	29.5	38.7	35.4	1.3	41.4	35.5	47.3	48.0	6.6
21	Puebla	19.4	15.9	22.9	25.3	5.9	23.2	18.2	28.2	29.0	5.8
22	Querétaro	9.9	7.6	12.1	12.1	2.2	13.5	9.8	17.3	17.0	3.4
23	Quintana Roo	21.6	18.9	24.3	25.8	4.3	5.3	3.6	7.1	14.3	9.0
24	San Luis Potosí	16.4	12.2	20.5	16.9	0.6	25.8	20.7	30.8	27.3	1.5
25	Sinaloa	8.4	6.5	10.2	13.8	5.4	11.7	8.4	15.1	14.3	2.6
26	Sonora	11.9	9.3	14.5	13.0	1.1	14.2	8.5	19.8	13.7	0.5
27	Tabasco	21.7	18.1	25.3	17.6	4.1	23.6	16.5	30.7	34.4	10.8
28	Tamaulipas	9.7	6.0	13.3	11.8	2.1	17.0	8.8	25.2	13.7	3.3
29	Tlaxcala	11.8	9.7	13.9	14.2	2.4	6.9	4.7	9.1	10.3	3.4
30	Veracruz	24.5	20.6	28.4	25.7	1.2	29.5	23.5	35.4	34.3	4.9
31	Yucatán	19.5	16.6	22.3	22.4	2.9	20.7	17.7	23.8	23.8	3.1
32	Zacatecas	5.8	3.8	7.9	8.4	2.5	16.1	11.4	20.8	14.8	1.3
Estados Unidos Mexicanos		15.2	14.4	16.0	17.0	1.7	16.5	15.5	17.4	19.3	2.8

Nota: Estimaciones del CONEVAL con base en la muestra del Censo de Población y Vivienda 2010 y el MCS-ENIGH 2010.

Cuadro C.4
Comparación de las estimaciones de población con ingreso inferior a la línea de bienestar, según entidad federativa. México, 2010.

Entidad	Población con ingreso inferior a la línea de bienestar			Error absoluto relativo ¹	
	MCS-ENIGH	Mejor Predictor	Elbers y	Mejor Predictor	Elbers y
		Empírico	colaboradores		
Aguascalientes	46.2	46.7	47.3	1.1	2.4
Baja California	38.2	35.1	40.5	8.0	6.2
Baja California Sur	35.4	32.9	35.4	7.0	0.0
Campeche	54.6	54.0	56.3	1.2	3.1
Coahuila	40.6	46.5	45.2	14.7	11.3
Colima	39.6	37.8	42.0	4.5	6.2
Chiapas	80.9	77.5	78.0	4.2	3.5
Chihuahua	51.8	46.9	45.7	9.5	11.9
Distrito Federal	34.0	33.5	35.6	1.2	4.7
Durango	60.1	55.5	54.8	7.7	8.9
Guanajuato	54.2	51.5	55.5	4.9	2.4
Guerrero	69.6	68.4	73.6	1.7	5.8
Hidalgo	58.8	56.5	58.4	4.0	0.7
Jalisco	43.0	43.2	45.5	0.3	5.9
México	48.4	49.9	51.6	2.9	6.6
Michoacán	59.0	56.9	62.8	3.5	6.5
Morelos	49.1	49.4	50.3	0.5	2.3
Nayarit	45.6	47.1	48.6	3.2	6.6
Nuevo León	29.3	35.7	36.0	22.2	23.1
Oaxaca	68.6	70.4	71.4	2.6	4.1
Puebla	66.7	64.5	66.4	3.3	0.5
Querétaro	46.3	44.0	46.3	5.0	0.1
Quintana Roo	39.2	40.3	49.5	2.8	26.4
San Luis Potosí	59.5	58.4	58.4	1.8	1.7
Sinaloa	44.1	46.5	46.9	5.4	6.2
Sonora	40.2	40.4	42.4	0.5	5.4
Tabasco	61.3	54.0	55.5	11.8	9.4
Tamaulipas	48.4	47.1	47.2	2.8	2.6
Tlaxcala	67.6	64.6	60.4	4.4	10.6
Veracruz	62.6	59.1	60.4	5.7	3.6
Yucatán	54.8	56.4	57.9	3.0	5.7
Zacatecas	67.0	61.3	58.8	8.5	12.3
Nacional	52.0	51.3	52.8	1.3	1.6

Fuente: estimaciones del CONEVAL con base en el MCS-ENIGH 2010 y la muestra del Censo de Población y Vivienda 2010.

¹ El error absoluto relativo es el cociente del valor absoluto de la diferencia de la estimación y el valor verdadero entre el valor verdadero.

Cuadro C.4 (continuación)
Comparación de las estimaciones de población con ingreso inferior a la línea de bienestar mínimo, según entidad federativa. México, 2010.

Entidad	Población con ingreso inferior a la línea de bienestar mínimo			Error absoluto relativo ¹	
	MCS-ENIGH	Mejor Predictor Empírico	Elbers y colaboradores	Mejor Predictor Empírico	Elbers y colaboradores
Aguascalientes	14.8	19.8	12.4	33.8	15.9
Baja California	9.8	11.9	9.9	21.3	0.5
Baja California Sur	11.2	11.1	8.1	0.3	27.6
Campeche	21.4	26.7	23.0	24.6	7.2
Coahuila	11.4	18.7	11.4	63.3	0.1
Colima	8.5	13.4	11.0	58.6	30.1
Chiapas	50.9	55.2	48.9	8.5	4.0
Chihuahua	16.6	20.5	13.7	23.4	17.3
Distrito Federal	6.0	11.7	7.8	94.7	30.3
Durango	23.1	28.6	19.6	23.5	15.2
Guanajuato	16.5	24.0	18.6	46.1	13.1
Guerrero	38.5	43.9	42.2	14.1	9.6
Hidalgo	23.7	29.4	24.5	23.8	3.4
Jalisco	14.4	17.5	12.5	21.5	13.3
México	14.4	21.5	16.2	49.3	12.4
Michoacán	21.5	28.8	26.2	33.8	21.9
Morelos	14.0	22.2	16.9	58.2	20.2
Nayarit	17.7	22.5	17.0	27.1	4.1
Nuevo León	6.0	12.6	7.7	109.0	26.7
Oaxaca	36.8	47.3	40.5	28.5	10.1
Puebla	27.3	36.3	32.9	33.1	20.5
Querétaro	16.0	19.2	14.6	19.8	8.9
Quintana Roo	11.9	16.6	17.1	39.8	43.7
San Luis Potosí	26.0	31.7	24.7	22.0	4.9
Sinaloa	14.1	20.0	13.5	42.0	3.9
Sonora	11.0	15.6	11.2	41.7	1.2
Tabasco	22.4	26.5	22.1	18.1	1.6
Tamaulipas	14.6	19.5	13.3	33.9	9.1
Tlaxcala	26.7	33.7	21.9	26.3	18.0
Veracruz	28.3	32.2	28.4	13.9	0.2
Yucatán	17.9	27.5	24.1	53.8	34.7
Zacatecas	29.5	33.5	22.0	13.7	25.6
Nacional	19.4	25.2	20.0	30.2	3.3

Fuente: estimaciones del CONEVAL con base en el MCS-ENIGH 2010 y la muestra del Censo de Población y Vivienda 2010.

¹ El error absoluto relativo es el cociente del valor absoluto de la diferencia de la estimación y el valor verdadero entre el valor verdadero.

Cuadro C.4 (continuación)
Promedio de los errores absolutos relativos,¹
según método de estimación, México, 2010

Indicadores	Mejor Predictor Empírico	Elbers y colaboradores
Población con ingreso inferior a la línea de bienestar	5.0	6.5
Población con ingreso inferior a la línea de bienestar mínimo	35.0	14.2
Ambos indicadores	20.0	10.3

Fuente: estimaciones del CONEVAL con base en el MCS-ENIGH 2010 y la muestra del Censo de Población y Vivienda 2010.

¹ El error absoluto relativo es el cociente del valor absoluto de la diferencia de la estimación y el valor verdadero entre el valor verdadero.

Cuadro C.5
Indicadores que se utilizaron para la construcción de
los estratos socioeconómicos de los municipios. México, 2010

Variables
Ingreso
Porcentaje de la población en situación de pobreza patrimonial (2005)
Rezago educativo
Porcentaje de la población de 3 a 15 años con rezago educativo
Porcentaje de la población de 16 años o más con rezago educativo
Acceso a los servicios de salud
Porcentaje de la población sin derecho-habienencia a servicios de salud
Proporción de los hijos fallecidos de las mujeres de 15 a 49 años
Acceso a la seguridad social
Porcentaje de la población ocupada (16 años o más) sin servicio médico ni sistema de ahorro para el retiro
Calidad y espacios de la vivienda
Porcentaje de la población en viviendas particulares habitadas con pisos de tierra
Porcentaje de la población en viviendas particulares habitadas con carencia en material de muros
Porcentaje de la población en viviendas particulares habitadas con carencia en material de techos
Porcentaje de la población en viviendas con hacinamiento
Servicios básicos en la vivienda
Porcentaje de la población en viviendas particulares habitadas con carencia en el acceso al agua entubada
Porcentaje de la población en viviendas particulares habitadas con carencia en drenaje
Porcentaje de la población en viviendas particulares habitadas con carencia en energía eléctrica
Porcentaje de la población en viviendas particulares habitadas con carencia por combustible para cocinar
Porcentaje de viviendas particulares habitadas que no disponen de excusado o sanitario
Porcentaje de viviendas sin servicio de eliminación de basura
Alimentación
Porcentaje de baja talla para edad en menores de 5 años (2005-2006)
Porcentaje de bajo peso para edad en menores de 5 años (2005-2006)
Grupo étnico
Porcentaje de la población que se considera indígena o habla una lengua indígena

Fuente: elaboración del CONEVAL con información de la ENIGH 2005, el Censo de Población y Vivienda 2005, los Mapas de Nutrición en México 2005-2006 y la Muestra del Censo de Población y Vivienda 2010.

Cuadro C.6
Medidas y pruebas estadísticas con cuatro clases latentes

Medida	Valor
BIC-Ajustado	338,084.3
Prueba Lo-Mendel Rubin, p-value	0.001
Entropía	0.9
Proporción de municipios en el estrato 1	16.6
Proporción de municipios en el estrato 2	36.2
Proporción de municipios en el estrato 3	33.5
Proporción de municipios en el estrato 4	13.7

Fuente: estimaciones del CONEVAL con base en la construcción de clases latentes a partir de las 19 variables de pobreza (cuadro C. 5), con el paquete M Plus.

Cuadro C.7
Media de los indicadores que se utilizaron para la construcción de
los estratos socioeconómicos de los municipios. México, 2010

Indicador	Estratos			
	1	2	3	4
Ingreso				
Porcentaje de la población en situación de pobreza patrimonial (2005)	91.4	81.9	61.6	38.2
Educación				
Porcentaje de la población de 3 a 15 años con rezago educativo	19.4	15.5	15.4	12.8
Porcentaje de la población de 16 años o más con rezago educativo	62.7	47.7	35.7	17.6
Salud				
Porcentaje de la población sin derecho-habienencia a servicios de salud	44.6	41.8	37.8	29.2
Proporción de los hijos fallecidos de las mujeres de 15 a 49 años	6.7	4.5	3.9	2.8
Seguridad social				
Porcentaje de la población ocupada (a partir de 16 años) sin servicio médico ni sistema de ahorro para el retiro	97.2	93.3	85.8	62.6
Calidad y espacios de la vivienda				
Porcentaje de la población en viviendas particulares habitadas con pisos de tierra	24.4	15.9	7.4	2.1
Porcentaje de la población en viviendas particulares habitadas con carencia en material de muros	9.9	5.4	2.0	1.1
Porcentaje de la población en viviendas particulares habitadas con carencia en material de techos	8.4	6.0	3.4	1.4
Porcentaje de la población en viviendas particulares habitadas con hacinamiento	35.6	22.7	14.1	7.3
Servicios básicos en la vivienda				
Porcentaje de la población en viviendas particulares habitadas con carencia en el acceso al agua entubada	40.1	28.3	15.5	5.2
Porcentaje de la población en viviendas particulares habitadas con carencia en drenaje	60.8	32.3	14.9	3.9
Porcentaje de la población en viviendas particulares habitadas con carencia en energía eléctrica	12.4	4.4	2.1	0.7
Porcentaje de la población en viviendas particulares habitadas con carencia por combustible para cocinar	53.2	27.9	7.1	0.7
Porcentaje de la población en viviendas particulares habitadas que no disponen de excusado o sanitario	17.7	10.9	6.7	1.3
Porcentaje de la población en viviendas particulares habitada sin servicio de eliminación de basura	13.7	12.0	6.0	1.7
Alimentación				
Porcentaje de baja talla para edad en menores de 5 años (2005-2006)	38.2	25.3	16.6	11.0
Porcentaje de bajo peso para edad en menores de 5 años (2005-2006)	19.1	14.4	10.1	7.8
Lengua indígena				
Porcentaje de la población que se considera indígena o habla una lengua indígena	77.8	42.2	12.8	7.5

Fuente: Estimaciones del CONEVAL con base en la Muestra del Censo de Población y Vivienda 2010, el conteo de Población y Vivienda 2005, la Encuesta Nacional de Ingreso y Gasto en los Hogares 2005, y los Mapas de Nutrición en México 2005-2006

Cuadro C.8

Porcentaje en los indicadores relacionados con el ingreso, en la medición de la pobreza, según estrato rural y urbano. México, 2010

Indicador de pobreza	Estratos							
	1		2		3		4	
	Urbano	Rural	Urbano	Rural	Urbano	Rural	Urbano	Rural
Población con un ingreso inferior a la línea de bienestar mínimo	53.2	66.9	35.2	42.8	22.4	23.5	10.7	13.9
Población con un ingreso inferior a la línea de bienestar	79.5	90.1	76.4	75.3	63.7	59.1	41.2	40.2

Fuente: estimaciones del CONEVAL con base en el MCS-ENIGH 2010.