



SEDESOL

SECRETARÍA DE
DESARROLLO SOCIAL

ANEXO B

MODELO ÚNICO DE CLASIFICACIÓN (MUC) FORMULACIÓN DEL MUC

**SUBSECRETARÍA DE PLANEACIÓN, EVALUACIÓN Y DESARROLLO REGIONAL
DIRECCIÓN GENERAL DE GEOESTADÍSTICA Y PADRONES DE BENEFICIARIOS**

Enero 2015

CONTENIDO

1. JUSTIFICACIÓN TEORICA DEL MUC	4
2. MODELO ÚNICO DE CLASIFICACIÓN (MUC)	5
3. ESTRATEGIA METODOLÓGICA.....	6
Definición de estratos.....	6
Definición de variables.....	6
Formación de grupos.....	7
Discriminación entre grupos.....	7
Caracterización de los grupos y selección de la agrupación.....	7
Creación de indicadores.....	8
4. DEFINICIÓN DE VARIABLES.....	8
Rezago educativo	9
Acceso a servicios de salud	10
Acceso a la seguridad social.....	10
Calidad y espacios en la vivienda	11
Acceso a servicios básicos en la vivienda.....	11
Acceso a la alimentación	11
Ingreso	12
5. FORMACIÓN DE GRUPOS	12
6. DISCRIMINACIÓN ENTRE GRUPOS	17
7. CARACTERIZACIÓN DE GRUPOS	28
8. SELECCIÓN DE LA AGRUPACIÓN.....	34
Agrupación 1*	34
Agrupación 2*	34
Agrupación 3*	35
Agrupación 4*	35

Agrupación Seleccionada	35
9. CREACIÓN DE INDICADORES	36
Indicador intra-grupo.	37
Indicadores intra-estrato y nacional.	39
10. VENTAJAS Y DESVENTAJAS DEL MUC	40

1. JUSTIFICACIÓN TEORICA DEL MUC

La Secretaría de Desarrollo Social (SEDESOL) a través de la Dirección General de Geoestadística y Padrones de Beneficiarios (DGGPB) requiere de una herramienta estadística, homogénea y no discrecional, que considere las dimensiones de ingreso, carencia e intensidad de las carencias, con la cual se clasifique a los hogares y personas con base en sus condiciones socioeconómicas y demográficas, y que le permita generar un criterio de priorización para la intervención de los programas y la realización de análisis longitudinales.

El gobierno de México a través de la SEDESOL ha diseñado una serie de programas sociales con la idea de mejorar las condiciones de bienestar de los hogares más necesitados. La focalización de los programas consiste en dirigir acciones a una población objetivo asegurándose que las acciones y apoyos lleguen a quien más lo necesita.

En el año 2009 el Consejo Nacional de Evaluación (CONEVAL) da a conocer una metodología multidimensional para la definición, identificación y medición de la pobreza y en 2010 establece en sus lineamientos que los programas deben identificar a las personas o grupos de personas en situación de pobreza usando los siguientes criterios (o combinación de los mismos).

- Criterios asociados al bienestar económico, resultado de comparar el ingreso mensual per cápita del hogar, con el valor de las líneas de bienestar (LB) o bienestar mínimo (LBM).
- Criterios asociados a las carencias sociales: rezago educativo, acceso a los servicios de salud, acceso a la seguridad social, calidad y espacios de la vivienda, servicios básicos en la vivienda y acceso a la alimentación.
- Criterios asociados al contexto territorial (infraestructura social básica, cohesión social, etc.).

En la medición multidimensional de la pobreza el CONEVAL define que una persona está en pobreza multidimensional si está por debajo de la línea de bienestar económico (ingreso inferior a la LB) y presenta una o más de las carencias sociales o de contexto territorial.

En la actualidad, los beneficiarios de los programas sociales de la SEDESOL son identificados utilizando la dimensión del ingreso, alguna

carencia social y en la mayoría de los casos criterios operativos definidos por cada programa.

La SEDESOL busca realizar un cambio en la focalización de los programas sociales y propone la creación de un Modelo Único de Clasificación (MUC) que determine una agrupación de hogares con características sociales, económicas y demográficas similares y a partir de esta clasificación generar criterios de elegibilidad de los hogares a los distintos programas sociales, considerando las dimensiones de pobreza multidimensional.

2. MODELO ÚNICO DE CLASIFICACIÓN (MUC)

El objetivo general que busca el presente estudio es:

Generar una metodología para la creación de grupos homogéneos de hogares de acuerdo con las condiciones socioeconómicas y demográficas de los hogares, e identificar una forma en la que un hogar nuevo en la muestra, pueda ser asignado o clasificado en uno de los grupos creados.

Características del modelo:

- El modelo debe establecer criterios de elegibilidad y de priorización para la intervención adecuada de programas en los hogares.
- La clasificación de los grupos debe ser clara y medible, de tal manera que la observación de las condiciones del hogar a lo largo del tiempo y a partir de intervenciones específicas, permitan cuantificar los cambios en las condiciones del mismo, es decir que permitan dar un seguimiento longitudinal.

3. ESTRATEGIA METODOLÓGICA

Definición de estratos.

El análisis de hogares se realiza para cada uno de los estratos, los cuales se determinan de acuerdo al número de habitantes en la localidad. Los estratos a utilizar son:

- Estrato 1 o metropolitano (E1), con localidades de más de 100,000 habitantes
- Estrato 2 o urbano (E2), con localidades mayores a 15,000 y menores a 100,000 habitantes.
- Estrato 3 o semiurbano (E3), con localidades mayores a 2,500 y menores a 15,000 habitantes
- Estrato 4 o rural (E4), con localidades menores a 2,500 habitantes.

Para la realización del estudio se utiliza la información del Módulo de Condiciones Socioeconómicas de la Encuesta Nacional de Ingresos y Gastos en los Hogares (MCS-ENIGH 2012).

A continuación se describen los pasos metodológicos realizados en el diseño del modelo: definición de las variables, creación de grupos, discriminación entre grupos, caracterización de grupos, selección de la agrupación y creación de indicadores.

Definición de variables.

Con la finalidad de encontrar las características que describan mejor las condiciones de los hogares de tal manera que pueda realizarse la discriminación entre grupos, se crearon variables en función de las carencias definidas por el CONEVAL que incluyen las dimensiones de bienestar económico y derechos sociales; específicamente las dimensiones de los derechos sociales (rezago educativo, acceso a servicios de salud, acceso a la seguridad social, calidad y espacios de la vivienda, acceso a servicios básicos de vivienda, acceso a la alimentación), el ingreso, que pertenece a la dimensión de bienestar económico, y el contexto territorial en el análisis por estrato.

Para mejorar la discriminación entre hogares, además de las variables de carencias se consideró utilizar mediciones de la intensidad en los indicadores de pobreza multidimensional que refleje la profundidad de la

carencia presentada en el hogar y sus integrantes. Adicionalmente se construyeron indicadores que reflejan carencias e intensidad por grupos de edad de los integrantes con lo cual se obtuvieron valores diferenciados conforme al indicador y composición demográfica del hogar, de esta manera el modelo es sensible a estas poblaciones vulnerables.

Formación de grupos.

Una vez definidas las variables a considerar en el análisis, en esta fase se procede a formar grupos de hogares en condiciones socioeconómicas similares dentro de cada estrato. Con ello se estará en condiciones de obtener grupos de hogares homogéneos que permitan la focalización de los programas sociales. Para ello se utilizarán distintas metodologías en la creación de grupos con la finalidad de tener distintas agrupaciones formadas con criterios diferentes que permitan tener varias alternativas y poder elegir aquella que presente la mejor solución para los objetivos del presente estudio, considerando las diferentes aristas de la pobreza.

Discriminación entre grupos.

Este proceso de discriminación entre grupos permite identificar las variables que son las más relevantes en la formación de los grupos. Este paso consiste en determinar un criterio de identificación de cada grupo generado. De esta manera se estará en condiciones de evaluar cuáles variables son más relevantes en la formación de grupos y permite conocer la eficiencia de la clasificación. Para ello se usarán técnicas de análisis discriminante. Adicional a la identificación de las variables más relevantes en la formación de los grupos, se reagrupan los hogares para que satisfagan el criterio de clasificación de tal manera que los grupos queden perfectamente identificados por la regla de clasificación y no se tengan errores en la asignación.

Caracterización de los grupos y selección de la agrupación.

Una vez que se obtiene la segmentación de los hogares en grupos, el cuarto paso consiste en estudiar los perfiles de los grupos usando variables socioeconómicas, además de las variables originalmente usadas en el

proceso de creación de los grupos. Este paso es importante porque permite seleccionar al tipo de agrupación que mejor se ajuste a los objetivos planteados de focalización.

Mediante el conocimiento de los perfiles de los grupos y conociendo las características de las distintas propuestas de agrupaciones, se selecciona aquella que cumpla lo más posible con la identificación de los grupos con distintas carencias de pobreza multidimensional.

Es decir, en este paso se usan tanto criterios estadísticos como sustantivos para elegir el mejor método de agrupación, los criterios estadísticos sugieren ciertas reglas de selección de los modelos y los criterios sustantivos usan variables asociadas a las reglas de operación de los programas. Ambos criterios se usan para elegir el mejor método de clasificación y permiten evaluar cuáles variables son las que ejercen mayor influencia en la caracterización de los grupos y de esta manera seleccionar los hogares potenciales de los programas sociales

Creación de indicadores.

Además de clasificar a los hogares en grupos de pobreza multidimensional por estrato, similar con la agrupación seleccionada en el paso anterior, se considera importante un quinto paso que consiste en la creación de un indicador que permita priorizar las intervenciones ya sea dentro del mismo grupo, comparando entre grupos del mismo estrato o comparando a nivel nacional.

En las Secciones 4 a 8 se describen con mayor detalle los pasos descritos con anterioridad.

4. DEFINICIÓN DE VARIABLES

El CONEVAL en su “Metodología para la medición multidimensional de la pobreza en México” (CONEVAL, 2009), establece los criterios para determinar las condiciones de pobreza de los hogares. En el documento se establecen dos dimensiones medibles de pobreza llamadas dimensiones de pobreza multidimensional: bienestar económico y derechos sociales. La dimensión del bienestar económico se mide exclusivamente por el ingreso, en cambio la dimensión de derechos sociales se mide a través de seis

indicadores que son: rezago educativo, acceso a servicios de salud, acceso a la seguridad social, calidad y espacios de la vivienda, acceso a servicios básicos de vivienda y acceso a la alimentación. El CONEVAL determina las carencias de personas pero el modelo único de clasificación incorpora dimensiones del hogar y personas, de alguna manera se “hereda” la carencia del hogar hacia las personas.

Con base en la definición de las variables de carencias del CONEVAL, se propone utilizar los mismos indicadores pero con variaciones que permitan, además de identificar si un hogar presenta una carencia específica, determinar el grado de profundidad en que el hogar presenta la carencia. Adicionalmente a la profundidad de la carencia, se considera importante separar algunos indicadores de carencias relativos a la población de menores de edad o de adultos.

Las variables seleccionadas a partir de la revisión de los avances en reuniones de trabajo con personal de la SEDESOL, son once indicadores de intensidad de las carencias que a continuación se describen.

Rezago educativo

Para el indicador de *rezago educativo* se proponen dos indicadores de intensidad divididos de acuerdo a la edad

- **rezedu15:** Porcentaje de miembros del hogar entre 3 y 15 años con rezago educativo.

Calculado de la siguiente manera:

$$\text{rezedu15} = \text{rezedu3a15} / \text{p3a15}$$

donde *rezedu3a15* es el número de miembros del hogar entre 3 y 15 años con rezago educativo, y *p3a15* es el número de miembros del hogar entre 3 y 15 años.

- **rezedu16m:** Porcentaje de miembros del hogar de 16 años o más con rezago educativo.

Calculado de la siguiente manera:

$$\text{rezedu16m} = \text{rezedu16omas} / \text{p16omas}$$

donde *rezedu16omas* es el número de miembros del hogar de 16 años o más con rezago educativo y *p16omas* es el número de miembros del hogar de 16 años o más.

Acceso a servicios de salud

Para el indicador de *acceso a servicios de salud* se propone un indicador ponderado en el hogar, de tal forma que le dé más intensidad a la carencia (100%) si el *i*-ésimo miembro del hogar no tienen acceso a servicios de salud ($serv_sal_i = 0$), un peso del 50% si el miembro del hogar tienen acceso al seguro popular ($serv_sal_i = 1$) y un peso de cero si el miembro del hogar tiene acceso a algún servicio de salud como IMSS, ISSSTE, Pemex o algún otro servicio médico ($serv_sal_i = 2,3,4,5$).¹ Si se considera un hogar con *m* miembros, el indicador ponderado toma valores en el intervalo (0,1) y se define como

- **asalud** = $\frac{1}{m} \sum_{i=1}^m 1 * I(serv_sal_i = 0) + 0.5 * I(serv_sal_i = 1) + 0 * I(serv_sal_i = 2,3,4,5)$.

Para entender mejor este indicador se puede tomar como ejemplo una familia con 4 miembros con los siguientes escenarios: i) Si ningún miembro tiene acceso a servicios de salud entonces el indicador es la unidad (asalud=1); ii) si todos los miembros del hogar tienen acceso al seguro popular entonces el indicador es 0.5 (asalud=0.5); iii) si dos miembros tienen acceso al IMSS y los otros 2 miembros tienen acceso al seguro popular, entonces el indicador es 0.25 (asalud=0.25); iv) si todos los miembros tienen acceso al IMSS entonces el indicador asigna un cero (asalud=0).

Acceso a la seguridad social

En cuanto al indicador de *seguridad social* se propone otro indicador ponderado, dando el mismo peso a todos los miembros del hogar. Este indicador se define como

- **segsoc**: Porcentaje de miembros del hogar con carencia de acceso a seguridad social.

Esta variable también se denota como *ic_segsocintpc* y se calcula de la siguiente manera:

$$ic_segsocintpct = ic_segsocinten / tamhog$$

donde *ic_segsocinten* es el número de miembros del hogar con carencia de acceso a seguridad social y *tamhog* es el número de miembros del hogar.

¹ Estos valores fueron obtenidos de la ENIGH

Calidad y espacios en la vivienda

En el indicador de *calidad y espacio de vivienda* se propone determinarlo mediante la proporción de los 4 aspectos de carencias que presenta el hogar. Este indicador toma valores en el intervalo (0,1) y se define como

- **cv**: Porcentaje de carencias entre piso, muros, techos y hacinamiento.

Esta variable también se denota como *ic_cvintpc* y se calcula de la siguiente manera

$$ic_cvintpct = ic_cvinten / 4$$

donde *ic_cvinten* es la suma de carencias que presenta el hogar entre pisos, muros, techos y hacinamiento.

Acceso a servicios básicos en la vivienda

En cuanto al indicador de acceso a *servicios básicos de vivienda*, al igual que el indicador anterior se propone un indicador que mida el porcentaje de carencias de los 4 servicios básicos de la vivienda. Este indicador se define como

- **sbv**: Porcentaje de carencias entre agua, luz, drenaje y combustible.²

Acceso a la alimentación

En la dimensión de acceso a la alimentación se propone generar dos indicadores de inseguridad alimentaria en una escala de 0 a 3, como lo define la variable de inseguridad alimentaria (*ins_ali*) propuesta por el CONEVAL, pero tomando en cuenta una división por edad: integrantes menores de 18 años e integrantes de 18 años o más. En el caso de que un hogar no tenga menores de edad el indicador de inseguridad alimentaria en menores toma el valor de cero. Estas dos variables toman valores en una escala de 0 a 3 en donde 0 indica sin carencia alimentaria, 1 inseguridad alimentaria leve, 2 inseguridad alimentaria moderada y 3 inseguridad alimentaria severa. Estos indicadores se definen como

- **ali17** = grado de inseguridad alimentaria en menores
- **ali18m** = grado de inseguridad alimentaria en adultos

² Los indicadores de agua, luz, drenaje, combustible, se obtienen de CONEVAL

Para la construcción de estas dos variables se usó la escala de inseguridad alimentaria del CONEVAL aplicada por separado a los menores de 18 años y a los de 18 años y más. Se considera sin inseguridad alimentaria a los miembros del grupo de edad si no se reportaron respuestas afirmativas; se consideran con inseguridad alimentaria leve a los miembros del grupo de edad si respondieron afirmativamente a un máximo de dos preguntas de la escala; se consideran con inseguridad alimentaria moderada a los miembros del grupo de edad si respondieron afirmativamente a tres o cuatro preguntas de la escala; se consideran con inseguridad alimentaria grave a los miembros del grupo de edad si respondieron afirmativamente a cinco o más preguntas de la escala.

Ingreso

En la dimensión de bienestar económico medida a través del ingreso se proponen utilizar tres indicadores basados en el ingreso per cápita sin considerar transferencias: el ingreso estimado en escala logarítmica y la probabilidad de que el ingreso real esté por debajo de las líneas de bienestar (LB) y bienestar mínimo (LBM). Estas variables son obtenidas con el modelo lineal sin considerar los factores de expansión y se denotan como:

- **ling** = Ingreso estimado en escala logarítmica.
- **pLB** = Probabilidad de que el ingreso real se encuentre debajo de la LB.
- **pLBM** = Probabilidad de que el ingreso real se encuentre por debajo de la LBM.

Para la definición de estas variables se sugiere ver el Anexo 3.

Estas once variables definen intensidades de carencias en las dos dimensiones de la pobreza multidimensional y sirvieron como base para la creación de agrupaciones por estrato con el cual se considera el contexto territorial.

5. FORMACIÓN DE GRUPOS

Para la generación de agrupaciones se usaron cuatro métodos de agrupamiento que permiten obtener grupos de diferente manera. Con esta

estrategia se puede comparar entre las distintas agrupaciones y seleccionar aquella que cumpla tanto con criterios estadísticos como sustantivos (explicados en la Sección 2) con las características requeridas por el MUC para lograr una mejor focalización.

Las primeras tres agrupaciones se basan en métodos de *agrupación no paramétricos* y la cuarta utiliza un *modelo paramétrico*; las cuatro estrategias se definen a continuación.

- La *metodología no paramétrica* a utilizar consiste en una combinación de métodos de agrupación jerárquicos y no jerárquicos en dos pasos. En un primer paso, dado el número tan grande de hogares a agrupar (un poco más de 55,000 hogares que representan a 31 millones al usar factores de expansión), se usa el método no jerárquico de k-medias con un número de grupos grande fijado en 30. En un segundo paso, considerando los centroides de los 30 grupos resultantes del paso uno como individuos representativos de cada grupo, se usa un método jerárquico de agrupación basado en distancias euclidianas y con liga completa, para reagrupar los 30 grupos iniciales en un número menor de grupos. El número final de grupos es el sugerido por el dendrograma (diagrama de árbol). Es importante notar que en esta metodología no paramétrica, tanto en la parte jerárquica como en la no jerárquica, todas las distancias son euclidianas, y en consecuencia, dependen de la escala de medición de las variables. Por lo tanto variables con escala mayor que otras dominarán las distancias. Es por esto que en las agrupaciones 2 y 3 eliminamos el efecto de la escala de medición al considerar algunas o todas las variables estandarizadas.
- La cuarta agrupación usa una *metodología paramétrica*. Específicamente se considera un modelo de clases latentes basado en mezclas de distribuciones normales multivariadas. Esta metodología es invariante ante transformaciones lineales de las variables por lo que usan las variables de intensidades en su escala original. El número de grupos se determina usando un criterio de selección de modelos.

Con la metodología no paramétrica, que combina métodos jerárquicos y no jerárquicos, se generaron tres agrupaciones distintas que se obtienen al considerar las variables de análisis en su escala original o estandarizada,

mientras que la cuarta agrupación se obtiene con la metodología paramétrica. Las cuatro agrupaciones se describen a continuación:

- Agrupación 1 (A1): Considera las 11 variables de intensidades de carencias en escala original, obteniendo 30 grupos con el método de k-medias y se reagrupan los centroides de cada grupo usando distancias euclidianas y el método jerárquico de liga completa.
- Agrupación 2 (A2): Considera la variable “ling”, que corresponde al ingreso estimado en escala logarítmica, estandarizada y las restantes 10 variables se utilizan en su escala original, obteniendo 30 grupos con el método de k-medias y se reagrupan los centroides de cada grupo usando distancias euclidianas y el método jerárquico de liga completa.
- Agrupación 3 (A3): Considera las 11 variables de intensidades de carencias estandarizadas, obteniendo 30 grupos con el método de k-medias y se reagrupan los centroides de cada grupo usando distancias euclidianas y el método jerárquico de liga completa.
- Agrupación 4 (A4): Considera las 11 variables en su escala original y genera los grupos usando la metodología de clases latentes.

A continuación se presentan algunos resultados importantes de las agrupaciones obtenidas. En las tablas, las columnas A1 a A4 denotan a cada una de las cuatro agrupaciones obtenidas por cada uno de los métodos y las filas E1 a E4 denotan cada estrato en los que se divide a los hogares.

Las agrupaciones A1 a la A3, se obtuvieron a partir de un método de agrupación jerárquico, por lo que producen dendrogramas (diagramas de árbol) en los cuales es posible identificar el número de grupos óptimo. En tanto, la agrupación A4 no requiere de dendrogramas.

Para entender mejor los diagramas, es necesario tomar en cuenta que entre más larga sean las líneas que unen un racimo con otro indica que los grupos que están siendo sostenidos por la misma rama son más heterogéneos entre sí (ver Figura 1). En cada diagrama se muestra con cuadros de color verde el corte que se hizo para identificar a los grupos en cada agrupación por estrato. En esta figura se muestran tres grupos formados (ver casilleros dibujados en verde) que conforman los grupos. El resto de los diagramas se presentan en las Figuras A1 a la A3 en el Anexo.

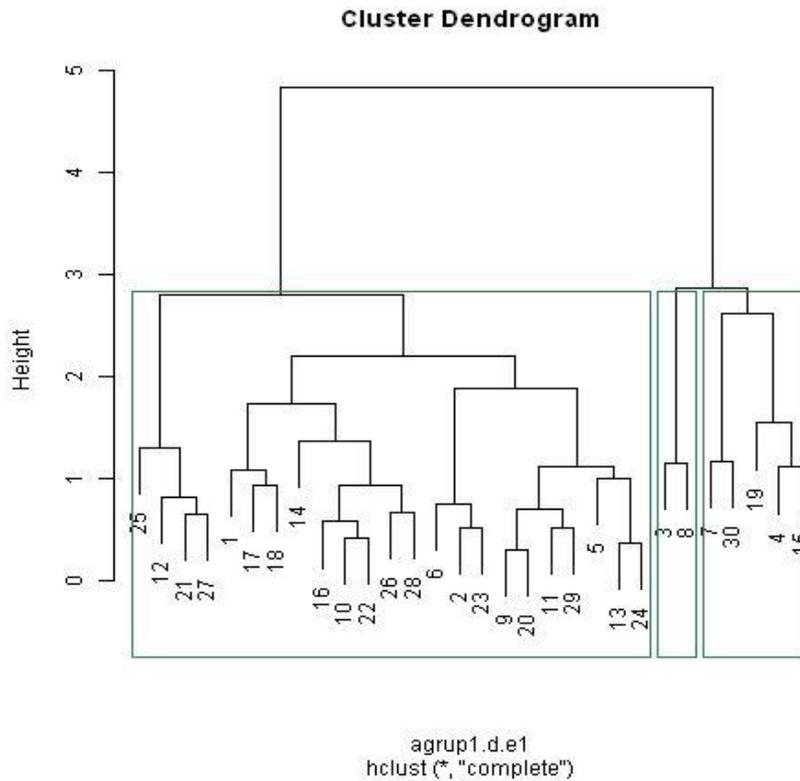


Figura 1. **Agrupación 1:** Dendrograma para el estrato 1

En la Tabla 1 se presenta el número de grupos, divididos por estrato, obtenidos por cada uno de los 4 métodos de agrupación. Se muestra también el total de grupos resultantes por método de agrupamiento obtenidos al sumar el número de grupos por columna. Por ejemplo, en la tabla se observa que la agrupación 3 es la que más grupos produce, 17 grupos en total con un mínimo de 4 grupos y un máximo de 5 por estrato. Por otro lado, las agrupaciones 2 y 4 son las que menos grupos tienen, 13 grupos en total con un mínimo de 3 y un máximo de 4 por estrato. La agrupación 1 es la que más variabilidad presenta con un mínimo de 3 grupos y un máximo de 5 grupos por estrato.

Tabla 1. Número de grupos en cada una de las 4 agrupaciones por estrato.

	Método de Agrupación			
Estrato	A1	A2	A3	A4
E1	3	3	4	3

E2	4	4	4	3
E3	5	3	4	3
E4	3	3	5	4
Total	15	13	17	13

En las tablas 2a a 2d se presenta la composición del número de hogares por estrato en cada una de las cuatro agrupaciones, respectivamente. Se puede observar que la agrupación 4 (Tabla 2d) es la que divide a los hogares en un número más homogéneo por estrato produciendo grupos de tamaño similar. Las otras tres agrupaciones (Tablas 2a, 2b y 2c) producen grupos de diversos tamaños, es decir, algunos grupos tienen muchos hogares mientras que otros tienen pocos.

Tabla 2a. **Agrupación 1:** Número de hogares por grupo y por estrato.

Estrato	Grupos ^a				
	G1	G2	G3	G4	G5
E1	13,012,197	568,075	1,898,496	NA	NA
E2	353,213	1,736,108	1,846,741	457,022	NA
E3	1,309,034	1,598,614	418,767	625,856	269,388
E4	2,430,042	1,570,902	2,580,816	NA	NA

(a) El valor de NA indica que no aplica

Tabla 2b. **Agrupación 2:** Número de hogares por grupo y por estrato.

Estrato	Grupos ^a			
	G1	G2	G3	G4
E1	9,350,646	5,332,285	795,837	NA
E2	1,264,742	1,650,543	1,003,283	474,516
E3	2,096,085	1,174,565	951,009	NA
E4	4,745,176	729,000	1,107,584	NA

(b) El valor de NA indica que no aplica

Tabla 2c. **Agrupación 3:** Número de hogares por grupo y por estrato.

Estrato	Grupos ^a				
	G1	G2	G3	G4	G5
E1	4,366,653	9,378,202	1,375,505	358,408	NA
E2	1,830,211	2,399,211	48,115	115,547	NA
E3	216,857	908,156	2,694,314	402,332	NA
E4	1,594,132	3,038,901	1,472,400	242,597	233,730

(a) El valor de NA indica que no aplica

Tabla 2d. **Agrupación 4: Número de hogares por grupo y por estrato.**

	Grupos ^a			
Estrato	G1	G2	G3	G4
E1	5,500,937	7,502,520	2,475,311	NA
E2	1,700,996	1,686,623	1,005,465	NA
E3	1,130,413	1,500,604	1,590,642	NA
E4	1,934,030	1,016,099	2,135,827	1,495,804

(a) El valor de NA indica que no aplica

6. DISCRIMINACIÓN ENTRE GRUPOS

En esta sección se identifican los criterios con los que cada grupo fue creado. Para ello se usan técnicas no paramétricas de análisis discriminante. En específico se usa la técnica de Árboles de Clasificación o comúnmente conocida como CART (*Classification and Regression Trees*, por sus siglas en inglés).

Para implementar cualquier análisis discriminante o de clasificación es necesario contar con una muestra de aprendizaje que en este caso son hogares previamente clasificados. La clasificación previa utilizada en esta fase de discriminación es la que describimos en la sección anterior (Formación de Grupos).

La técnica de árboles de clasificación utiliza una serie de biparticiones sucesivas del conjunto total de hogares de tal manera, que los grupos partidos (nodos hijos) sean más puros que el nodo padre en el sentido de que contengan hogares de la misma categoría. La técnica identifica las variables con mayor capacidad discriminatoria y determina ciertos rangos de valores de las variables con las que se forman las biparticiones. El algoritmo deja de partir los grupos cuando satisface cierto criterio de pureza máximo de cada grupo resultante. Los nodos que ya no se parten se llaman nodos terminales y a cada nodo terminal se le asigna una etiqueta de grupo. Para fines predictivos todos aquellos hogares que satisfacen los criterios para estar en cierto nodo terminal reciben el mismo valor del grupo asignado. Muchas veces como el algoritmo utiliza una búsqueda exhaustiva de manera computacional, es posible que un grupo que ya era lo

suficientemente puro sea partido. Por tal razón se opta por “podar” los árboles con criterios de optimización.

Se implementa entonces el algoritmo de árbol de clasificación utilizando como grupos conocidos los generados con cada una de las 4 agrupaciones (A1, A2, A3 o A4). Para cada estrato de cada agrupación se construye un árbol de clasificación. En total se generaron 16 árboles de clasificación. Las variables utilizadas para construir los árboles son las 11 variables de intensidades de carencias en su escala original. Transformaciones monótonas en las variables, como la estandarización o la función exponencial, no modifican el poder discriminatorio de los árboles y el usar las variables en su escala original simplifican la interpretación de los resultados.

Para tener una idea de la forma de los árboles se considera el árbol construido para el Estrato 1 de la Agrupación 1 que se presenta en la Figura 2. En esta figura se observan dos árboles que corresponden a la misma agrupación y al mismo estrato, la diferencia radica en que el árbol de la parte superior está un poco más crecido debido a que contiene más bifurcaciones. En cambio, el árbol de la parte inferior ha sido podado. La poda permite definir una clasificación más sencilla aunque aumenta el error de clasificación.

La forma en que debemos leer los árboles es de la siguiente manera. Si se toma el árbol de la Figura 2 en su parte inferior, si un hogar satisface el criterio de $\text{alim18} < 1.5$ se elige la rama izquierda, si además satisface el criterio de $\text{ali17} < 0.5$ entonces se elige nuevamente la rama izquierda y de esta manera se llega a un nodo terminal en donde ese hogar se asignará al grupo 1. En cambio, si un hogar satisface $\text{alim18} < 1.5$ (elegimos rama izquierda) pero $\text{ali17} > 0.5$ (elegimos derecha) entonces ese hogar se clasifica en el grupo 3. El resto de los árboles se muestran el Anexo 1 de Dendrogramas.

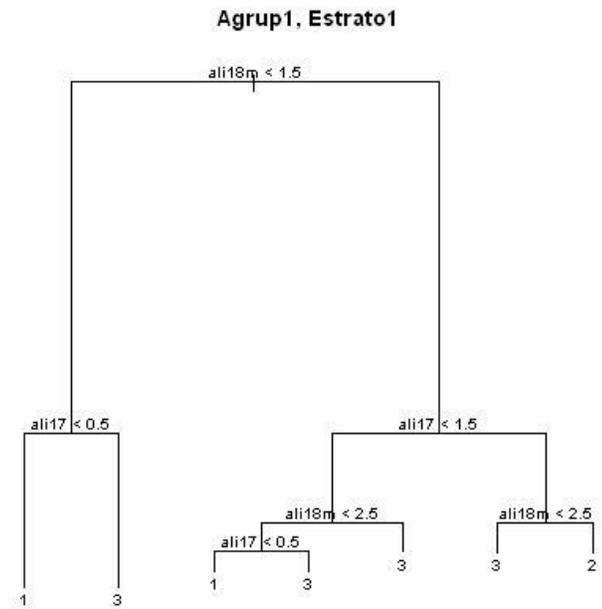
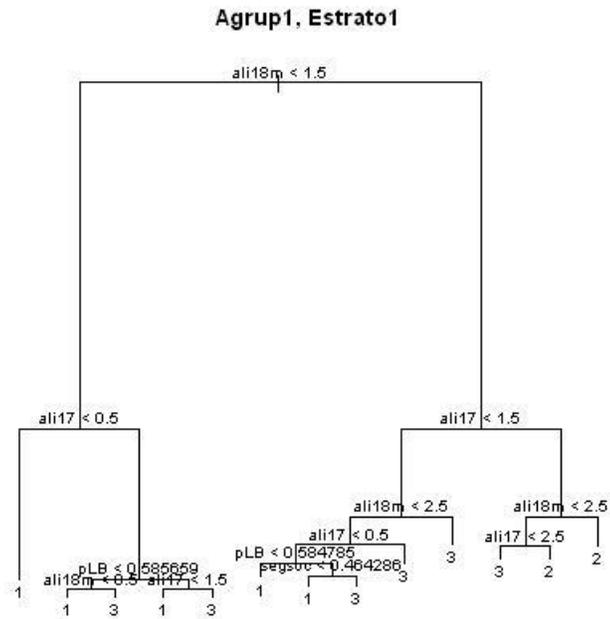


Figura 2. Ejemplo de árbol de clasificación para la Agrupación 1 y Estrato 1. Árbol sin podar (arriba) y árbol podado (abajo).

De esta manera se pueden conocer las variables relevantes para la clasificación de grupos para cada estrato. En total hay cuatro estratos y cuatro métodos que generan 16 árboles de clasificación.

Una forma de evaluar la eficiencia del método de discriminación es observar la tasa de errores de clasificación. Para cada uno de los 16 árboles se calcula el error de clasificación, como un indicador de qué tan bueno es el árbol para reproducir los grupos originales, y el número de nodos terminales (Nodos), como un indicador de la complejidad de la clasificación. La Tabla 3a presenta los errores de clasificación y el número de nodos terminales en un árbol sin podar.

Tabla 3a. **Árboles originales sin podar.** Tasa de errores de clasificación y número de nodos terminales en paréntesis (Nodos).

Estrato	Tasa de Errores Según Método de Agrupación							
	A1	Nodos	A2	Nodos	A3	Nodos	A4	Nodos
E1	0.008	(13)	0.042	(15)	0.062	(14)	0.038	(6)
E2	0.052	(11)	0.070	(12)	0.084	(16)	0.038	(7)
E3	0.065	(12)	0.045	(6)	0.046	(14)	0.039	(8)
E4	0.071	(10)	0.040	(14)	0.118	(16)	0.028	(5)
Total	0.041		0.046		0.076		0.036	

En la Tabla 3b se presenta la misma información que en la tabla anterior pero después de haber podado los árboles. En esta tabla se puede ver que los errores de clasificación de los árboles sin podar son menores que los errores de los árboles podados. En cambio el número de nodos terminales es mayor en los árboles originales. Al podar se disminuye la complejidad del árbol haciendo más fácil la clasificación, pero se incrementa el error de clasificación.

Los errores de clasificación para cada una de las 4 agrupaciones no son tan relevantes,³ pero son un indicador de la facilidad para reproducir la agrupación original con las variables de intensidades y con la metodología de los árboles de clasificación. En este sentido podemos notar que la Agrupación 3 es la más difícil de reproducir debido a que presenta los

³ Aunque sí sugieren cuales métodos de agrupación presentan una tasa de errores menor

errores de clasificación más altos y la Agrupación 4 es la más sencilla de reproducir porque presenta los errores de clasificación más bajos.

Tabla 3b. **Árboles podados.** Tasa de errores de clasificación y número de nodos terminales en paréntesis (Nodos).

Estrato	Tasa de Errores Según Método de Agrupación							
	A1	Nodos	A2	Nodos	A3	Nodos	A4	Nodos
E1	0.037	(7)	0.091	(7)	0.107	(6)	0.038	(3)
E2	0.090	(5)	0.095	(8)	0.113	(8)	0.049	(3)
E3	0.077	(7)	0.045	(3)	0.110	(7)	0.046	(3)
E4	0.083	(6)	0.092	(6)	0.142	(10)	0.028	(4)
Total	0.064		0.084		0.117		0.039	

Adicionalmente se tiene que la Agrupación 3 es la más compleja de todas debido a que tiene un número mayor de nodos terminales, seguida por la Agrupación 2, luego la 1 y finalmente la 4, siendo la más simple en cuanto a su especificación con el menor número de nodos terminales.

Para fines de este proyecto, se dio mayor importancia a los criterios discriminatorios sugeridos por los árboles de clasificación que a la agrupación original, por lo cual se reasignaron aquellos hogares pertenecientes a un grupo en la agrupación original y que son asignados a un grupo distinto por el árbol de clasificación.⁴ La clasificación asignada a cada hogar por los árboles de clasificación se denotan como Agrupación 1* a 4*.

Vale la pena notar que las agrupaciones originales y las resultantes de los árboles de clasificación no varían mucho, lo cual se puede confirmar por los errores de clasificación de 3.6% a 7.6% sin podar y de 3.9% a 11.7% después de la poda. Las nuevas distribuciones de hogares por estrato y por grupo en cada una de las cuatro Agrupaciones^{***} se muestran en las Tablas 4a a 4d. Comparando las Tablas 2a a 2d con las Tablas 4a a 4d se verifica que los tamaños de grupos presentan una variación mínima al reasignar los hogares con los árboles de clasificación.

⁴ En cada agrupación se reasignaron los casos después de la poda para mejorar la clasificación.

Tabla 4a. **Agrupación 1***: Número de hogares por grupo y por estrato.

Estrato	Grupos ^a				
	G1	G2	G3	G4	G5
E1	12,859,183	511,775	2,107,810	NA	NA
E2	369,150	1,615,718	1,827,810	580,406	NA
E3	1,151,100	1,704,236	416,247	751,783	198,293
E4	2,151,113	1,817,411	2,613,236	NA	NA

(a) El valor de NA indica que no aplica

Tabla 4b. **Agrupación 2***: Número de hogares por grupo y por estrato.

Estrato	Grupos ^a			
	G1	G2	G3	G4
E1	10,173,009	4,451,451	854,308	NA
E2	1,251,450	1,813,655	892,075	435,904
E3	2,070,652	1,168,786	982,221	NA
E4	4,600,949	652,680	1,328,131	NA

(a) El valor de NA indica que no aplica

Tabla 4c. **Agrupación 3***: Número de hogares por grupo y por estrato.

Estrato	Grupos ^a				
	G1	G2	G3	G4	G5
E1	3,586,699	10,265,499	1,509,092	117,478	NA
E2	2,204,266	2,008,641	67,650	112,527	NA
E3	175,626	690,596	2,712,283	643,154	NA
E4	1,529,999	3,125,252	1,731,382	55,349	139,778

(a) El valor de NA indica que no aplica

Tabla 4d. **Agrupación 4***: Número de hogares por grupo y por estrato.

Estrato	Grupos ^a			
	G1	G2	G3	G4
E1	5,435,226	7,665,794	2,377,748	NA
E2	1,706,134	1,739,652	947,298	NA
E3	1,081,338	1,489,282	1,651,039	NA
E4	1,912,932	1,024,368	2,134,504	1,509,956

(a) El valor de NA indica que no aplica

Adicionalmente, en las Tablas 5a a 5d, se presentan los tamaños de los grupos con respecto al número de individuos para cada una de las cuatro Agrupaciones “*”. Se observa que los tamaños de grupos por estrato en cada agrupación, ya sea en término de número de hogares (Tablas 4a a 4d) o en número de individuos (Tablas 5a a 5d) son los tamaños finales que se usan para comparar entre las distintas agrupaciones.

Tabla 5a. **Agrupación 1***: Número de individuos por grupo y por estrato.

	Grupos ^a				
Estrato	G1	G2	G3	G4	G5
E1	44,116,533	2,425,058	8,947,110	NA	NA
E2	1,802,948	6,595,721	6,359,404	2,033,312	NA
E3	4,594,889	6,008,101	1,147,721	3,768,778	1,027,821
E4	7,978,333	7,534,946	10,828,527	NA	NA

(a) El valor de NA indica que no aplica

Tabla 5b. **Agrupación 2***: Número de individuos por grupo y por estrato.

	Grupos ^a			
Estratos	G1	G2	G3	G4
E1	33,587,556	17,686,099	4,215,046	NA
E2	4,546,028	6,668,071	3,443,570	2,133,716
E3	8,389,994	4,182,483	3,974,833	NA
E4	18,086,187	3,278,358	4,977,261	NA

(a) El valor de NA indica que no aplica

Tabla 5c. **Agrupación 3***: Número de individuos por grupo y por estrato.

	Grupos ^a				
Estratos	G1	G2	G3	G4	G5
E1	14,052,715	33,735,547	7,202,449	497,990	NA
E2	8,663,227	7,356,064	286,229	485,865	NA
E3	848,618	2,721,249	10,007,130	2,970,313	NA
E4	6,762,418	12,054,226	6,551,331	258,380	715,451

(a) El valor de NA indica que no aplica

Tabla 5d. **Agrupación 4***: Número de individuos por grupo y por estrato.

Estratos	Grupos ^a			
	G1	G2	G3	G4
E1	20,073,408	24,178,694	11,236,599	NA
E2	6,431,083	6,030,285	4,330,017	NA
E3	3,913,694	6,722,771	5,910,845	NA
E4	9,184,195	3,972,720	7,794,850	5,390,041

(a) El valor de NA indica que no aplica

Las Figuras A5 a la A8 del Anexo muestran los 16 árboles de clasificación podados, los cuales muestran los criterios con los que se forman las nuevas Agrupaciones 1* a la 4*, respectivamente. Al comparar los árboles se puede notar que no se utilizan todas las variables para su construcción, lo que quiere decir que solo aquellas variables que aparecen en el árbol son necesarias para discriminar a los hogares en los distintos grupos.

Para tener una idea clara de la clasificación obtenida con las 4 Agrupaciones “*”, a continuación se resumen los criterios de pertenencia a cada grupo. Enseguida se presentan las condiciones en un mismo renglón separadas por comas lo cual significa que las condiciones se deben de satisfacer de manera simultánea, en cambio si la especificación de un mismo grupo tiene dos o tres líneas, las condiciones de cada línea son condiciones excluyentes (adicionales) que especifican la pertenencia al grupo. Por ejemplo, en la Agrupación 1* en el Estrato 1 y Grupo 3, los hogares que cumplan con las condiciones $ali18 < 1.5$ y $ali17 > 0.5$, ó $1.5 < ali18m < 2.5$ y $ali17 > 0.5$, ó $ali18m > 2.5$ y $ali17 < 1.5$ están dentro de este grupo.

Agrupación 1*:

- Estrato 1:
 - Grupo 1: $ali18m < 2.5$, $ali17 < 0.5$
 - Grupo 2: $ali18m > 2.5$, $ali17 > 1.5$
 - Grupo 3: $ali18m < 1.5$, $ali17 > 0.5$
 $1.5 < ali18m < 2.5$, $ali17 > 0.5$
 $ali18m > 2.5$, $ali17 < 1.5$
- Estrato 2:
 - Grupo 1: $ali18m > 1.5$, $ali17 > 1.5$
 - Grupo 2: $0.5 < ali18m < 1.5$, $pLB < 0.58$
 $ali18m < 1.5$, $pLB > 0.58$
 - Grupo 3: $ali18m < 0.5$, $pLB < 0.58$
 - Grupo 4: $ali18m > 1.5$, $ali17 < 1.5$

- Estrato 3:
 - Grupo 1: ali18m < 1.5, pLBM > 0.27, ali17 < 0.5, segsoc > 0.47
 - Grupo 2: ali18m < 1.5, pLBM < 0.27
ali18m < 1.5, pLBM > 0.27, ali17 < 0.5, segsoc < 0.47
 - Grupo 3: ali18m > 1.5, ali17 < 0.5
 - Grupo 4: ali18m < 1.5, pLBM > 0.27, ali17 > 0.5
ali18m > 1.5, 0.5 < ali17 < 2.5
 - Grupo 5: ali18m > 1.5, ali17 > 2.5
- Estrato 4:
 - Grupo 1: ali18m < 1.5, pLBM < 0.23
0.5 < ali18m < 1.5, 0.23 < pLBM < 0.60
 - Grupo 2: ali18m > 1.5
0.5 < ali18m < 1.5, pLBM > 0.60, ali17 > 1.5
 - Grupo 3: ali18m < 0.5, pLBM > 0.23
0.5 < ali18m < 1.5, pLBM > 0.60, ali17 < 1.5

En esta agrupación las variables más importantes para identificar los grupos son las intensidades de carencias de alimentación en personas mayores y menores de edad (ali18m y ali17) usadas en todos los estratos. Adicionalmente se usan las variables de ingreso que corresponden a la probabilidad de que el ingreso observado esté por debajo de la línea bienestar (pLB) y de bienestar mínimo (pLBM). También se usa la variable de intensidad de carencia de seguridad social (segsoc) en uno de los estratos.

Agrupación 2*:

- Estrato 1:
 - Grupo 1: pLBM < 0.15, ali18m < 0.5
pLBM < 0.15, 0.5 < ali18m < 1.5
pLBM > 0.15, ali17 < 1.5, segsoc < 0.61, ing > 1828.81
 - Grupo 2: pLBM < 0.15, ali18m > 1.5
pLBM > 0.15, ali17 < 1.5, segsoc < 0.61, ing < 1828.81
pLBM > 0.15, ali17 < 1.5, segsoc > 0.61
 - Grupo 3: pLBM > 0.15, ali17 > 1.5
- Estrato 2:
 - Grupo 1: ali18m < 0.5, asalud < 0.42, rezedu16m > 0.35
asalud > 0.42, segsoc < 0.85, ing < 2430.78
ali18m < 0.5, asalud > 0.42, segsoc > 0.85
 - Grupo 2: ali18m < 0.5, asalud < 0.42, rezedu16m < 0.35
asalud > 0.42, segsoc < 0.85, ing > 2430.78
ali18m > 0.5, ali17 < 1.5, pLBM < 0.22
 - Grupo 3: ali18m > 0.5, ali17 < 1.5, pLBM > 0.22
 - Grupo 4: ali18m > 0.5, ali17 > 1.5
- Estrato 3:
 - Grupo 1: ali18m < 1.5, pLB > 0.51
 - Grupo 2: ali18m < 1.5, pLB < 0.51
 - Grupo 3: ali18m > 1.5

- Estrato 4:
 - Grupo 1: $pLB < 0.48$, $ali18m > 1.5$
 $pLB > 0.48$, $ali17 < 1.5$
 $pLB > 0.48$, $ali17 > 1.5$, $ali18m < 1.5$
 - Grupo 2: $pLB > 0.48$, $ali17 > 1.5$, $ali18m > 1.5$
 - Grupo 3: $pLB < 0.48$, $ali18m < 1.5$

En esta agrupación las variables más importantes para discriminar entre los grupos son la probabilidad de que el ingreso observado esté por debajo de la línea de bienestar mínimo (pLBM), en los Estratos 1 y 4, y la variable de carencia de alimentación en adultos (ali18m) en los Estratos 2 y 3. Adicionalmente se usan las variables de carencia de alimentación en menores (ali17), el ingreso estimado (ing) y las probabilidades de estar por debajo de las líneas de bienestar y bienestar mínimo (pLB y pLBM). Se usan también las variables de seguridad social (segsoc), de carencias de acceso a servicios de salud (asalud) y una variable de rezago educativo en adultos (rezedu16m).

Agrupación 3*:

- Estrato 1:
 - Grupo 1: $pLBM > 0.13$, $ali17 < 1.5$, $cv < 0.13$, $asalud > 0.29$
 - Grupo 2: $pLBM < 0.13$, $rezedu15 < 0.58$
 $pLBM > 0.13$, $ali17 < 1.5$, $cv < 0.13$, $asalud < 0.29$
 - Grupo 3: $pLBM > 0.13$, $ali17 < 1.5$, $cv > 0.13$
 $pLBM > 0.13$, $ali17 > 1.5$
 - Grupo 4: $pLBM < 0.13$, $rezedu15 > 0.58$
- Estrato 2:
 - Grupo 1: $pLB < 0.65$, $rezedu16m < 0.48$, $rezedu15 < 0.75$, $ali18m < 1.5$, $cv > 0.13$
 $pLB < 0.65$, $rezedu16m < 0.48$, $rezedu15 < 0.75$, $ali18m > 1.5$
 $pLB < 0.65$, $rezedu16m > 0.48$
 $pLB > 0.65$, $rezedu15 < 0.58$, $cv < 0.38$
 - Grupo 2: $pLB < 0.65$, $rezedu16m < 0.48$, $rezedu15 < 0.75$, $ali18m < 1.5$, $cv < 0.13$
 - Grupo 3: $pLB > 0.65$, $rezedu15 < 0.58$, $cv > 0.38$
 - Grupo 4: $pLB < 0.65$, $rezedu16m < 0.48$, $rezedu15 > 0.75$
 $pLB > 0.65$, $rezedu15 > 0.58$
- Estrato 3:
 - Grupo 1: $cv < 0.13$, $rezedu15 > 0.37$
 - Grupo 2: $cv < 0.13$, $rezedu15 < 0.37$, $pLBM > 0.43$, $sbv < 0.13$, $ali18m > 1.5$
 $cv < 0.13$, $rezedu15 < 0.37$, $pLBM > 0.43$, $sbv > 0.13$
 - Grupo 3: $cv < 0.13$, $rezedu15 < 0.37$, $pLBM < 0.43$
 $cv < 0.13$, $rezedu15 < 0.37$, $pLBM > 0.43$, $sbv < 0.13$, $ali18m < 1.5$
 - Grupo 4: $cv > 0.13$
- Estrato 4:
 - Grupo 1: $pLBM > 0.46$, $sbv < 0.38$, $ali17 > 1.5$
 $pLBM > 0.46$, $sbv > 0.38$, $cv < 0.38$, $rezedu15 < 0.55$

- Grupo 2: $0.15 < pLBM < 0.46$, $segsoc > 0.41$, $ali17 < 1.5$
 $pLBM > 0.46$, $sbv < 0.38$, $ali17 < 1.5$
- Grupo 3: $pLBM < 0.15$
 $0.15 < pLBM < 0.46$, $segsoc < 0.41$
 $0.15 < pLBM < 0.46$, $segsoc > 0.41$, $ali17 > 1.5$
- Grupo 4: $pLBM > 0.46$, $sbv > 0.38$, $cv < 0.38$, $rezedu15 > 0.55$
- Grupo 5: $pLBM > 0.46$, $sbv > 0.38$, $cv > 0.38$

Esta agrupación es la más elaborada de todas en el sentido que casi todas las variables de intensidades de pobreza multidimensional (10 de 11 variables) son usadas para formar los grupos. Las variables más importantes varían de estrato a estrato y fueron las probabilidades de que el ingreso observado esté por debajo de las líneas de bienestar (pLB) y de bienestar mínimo (pLBM), en los Estratos 1,2 y 4, y la variable de calidad y espacios de la vivienda (cv) en el Estrato 3. Adicionalmente se usan las variables de alimentación (ali17 y ali18m), de rezago educativo (rezedu15 y rezedu16m), de seguridad social (segsoc), de acceso a servicios básicos de vivienda (sbv) y de acceso a servicios de salud (asalud). La única variable que no se usó para definir las agrupaciones es el ingreso estimado (ing).

Agrupación 4*:

- Estrato 1:
 - Grupo 1: $0.10 < pLBM < 0.35$
 - Grupo 2: $pLBM < 0.10$
 - Grupo 3: $pLBM > 0.35$
- Estrato 2:
 - Grupo 1: $pLBM > 0.15$, $ing > 1278.61$
 - Grupo 2: $pLBM < 0.15$
 - Grupo 3: $pLBM > 0.15$, $ing < 1278.61$
- Estrato 3:
 - Grupo 1: $pLBM < 0.49$, $ing > 2509.15$
 - Grupo 2: $pLBM > 0.49$
 - Grupo 3: $pLBM < 0.49$, $ing < 2509.15$
- Estrato 4:
 - Grupo 1: $pLB > 0.66$, $ing < 548.609$
 - Grupo 2: $pLB < 0.36$
 - Grupo 3: $pLB > 0.66$, $ing > 548.609$
 - Grupo 4: $0.36 < pLB < 0.66$

Esta agrupación es la más simple de todas en el sentido que todos los grupos quedan determinados únicamente por las tres variables de ingreso (ing, pLB y pLBM).

7. CARACTERIZACIÓN DE GRUPOS

Una vez definidos los grupos y después de establecer los criterios discriminatorios que los caracterizan, se analizan las características socioeconómicas de los hogares pertenecientes a cada grupo. La lista de variables a utilizar contiene 97 indicadores socioeconómicos. Debido al número tan grande de variables en el análisis, se decide conservar las 23 más relevantes para la focalización, que caracterizan los niveles de pobreza de un hogar. Para una mejor comparación entre las distintas agrupaciones, se presenta en las Tablas 6a a 6d, para los Estratos 1 al 4 respectivamente, una tabla comparativa de las cuatro Agrupaciones* por estrato de los promedios ponderados por grupo de estas 23 variables. La lista completa de variables se incorpora en archivos electrónicos que se entregan junto con este documento.

Al analizar las cuatro Tablas 6a a 6d se concluye lo siguiente:

Las Agrupaciones 1* y 4*, independientemente del estrato específico, forman grupos ordenados por un nivel global de pobreza multidimensional dominadas por la variable de ingreso (ing e ictSTpc).⁵ Por ejemplo, en el Estrato 1 (Tabla 6a), la Agrupación 1* coloca como grupo de pobreza global más intensa al grupo 2 (grupo de menor ingreso), seguido del 3 (ingreso medio) y por último el grupo 1 (ingreso más alto). Por su parte, la Agrupación 4*, en el mismo Estrato 1, coloca como grupo de pobreza más intensa al grupo 3, seguido del 1 y por último el 2. En estas dos agrupaciones, 1* y 4*, en las demás variables de pobreza, se mantiene el mismo orden indicado anteriormente.

La Agrupación 2* forma grupos ordenados por nivel global de pobreza multidimensional excepto en carencias de salud y seguridad social. Es decir, si se observa a las 11 variables que son utilizadas para definir los grupos, excepto las variables de asaludp y segsoc, los grupos obtenidos mantienen el mismo orden en esas variables. Por ejemplo, en el Estrato 2 (Tabla 6b), el grupo 4 es el que tiene un ingreso menor, tanto estimado como observado y también tiene las carencias mayores en calidad y espacios de la vivienda (cv), en acceso a servicios básicos de vivienda (sbv) y en rezago educativo en menores (rezedu15). Sin embargo, no es el grupo con promedios más

⁵ La variable ing se refiere a ingreso e ictSTpc se refiere a ingreso sin transferencias per cápita

altos en carencia de acceso a servicios de salud (asaludp) ni en carencia de acceso a la seguridad social (segsoc).

Tabla 6a. **Estrato 1:** Algunos perfiles socioeconómicos de las cuatro Agrupaciones*.

	Agrupación 1*			Agrupación 2*			Agrupación 3*				Agrupación 4*		
	G1	G2	G3	G1	G2	G3	G1	G2	G3	G4	G1	G2	G3
ictSTpc	5461.66	1988.16	2328.71	6342.34	2286.27	1709.68	2252.11	6301.81	1817.64	5504.27	2833.23	7445.07	1550.56
ing	4168.10	1530.64	2123.66	4852.31	1872.72	1356.66	1795.20	4844.96	1451.01	4200.22	2348.66	5660.72	1134.89
pLB	0.34	0.72	0.60	0.25	0.64	0.75	0.64	0.26	0.72	0.27	0.51	0.18	0.81
pLBM	0.13	0.43	0.32	0.07	0.33	0.46	0.32	0.08	0.43	0.07	0.20	0.04	0.51
asaludp	0.31	0.46	0.43	0.22	0.56	0.48	0.65	0.20	0.50	0.25	0.41	0.21	0.57
rezedu15	0.03	0.11	0.06	0.02	0.06	0.10	0.05	0.01	0.11	1.00	0.04	0.01	0.09
rezedu16m	0.13	0.29	0.22	0.09	0.25	0.28	0.24	0.09	0.30	0.06	0.20	0.05	0.33
ali17	0.00	2.60	1.03	0.05	0.21	2.40	0.14	0.07	1.50	0.12	0.24	0.05	0.76
ali18m	0.21	3.00	1.86	0.18	0.99	2.36	0.70	0.29	1.77	0.24	0.64	0.20	1.33
segsoc	0.44	0.66	0.60	0.32	0.77	0.69	0.84	0.31	0.70	0.35	0.53	0.33	0.81
cv	0.01	0.09	0.04	0.01	0.04	0.08	0.00	0.00	0.17	0.01	0.02	0.00	0.07
sbv	0.01	0.04	0.03	0.00	0.02	0.04	0.01	0.00	0.05	0.00	0.01	0.00	0.03
ic_rezedu	0.26	0.54	0.45	0.19	0.49	0.56	0.48	0.18	0.59	1.00	0.39	0.12	0.64
ic_asalud	0.35	0.42	0.41	0.29	0.52	0.44	0.61	0.26	0.45	0.34	0.42	0.27	0.50
ic_segsoc	0.60	0.81	0.74	0.50	0.88	0.84	0.95	0.49	0.83	0.56	0.70	0.49	0.91
ic_cv	0.04	0.25	0.13	0.02	0.12	0.24	0.00	0.01	0.55	0.04	0.06	0.01	0.22
ic_sbv	0.02	0.12	0.08	0.01	0.06	0.11	0.05	0.01	0.13	0.01	0.03	0.01	0.10
ic_ali	0.04	1.00	0.72	0.03	0.31	0.99	0.19	0.07	0.68	0.06	0.19	0.05	0.46
pobreza_e	0.01	0.17	0.10	0.00	0.06	0.17	0.05	0.00	0.16	0.00	0.02	0.00	0.14
pobreza_m	0.20	0.56	0.48	0.09	0.54	0.62	0.54	0.10	0.60	0.08	0.35	0.05	0.67
vul_car	0.33	0.27	0.32	0.33	0.33	0.20	0.33	0.34	0.24	0.53	0.37	0.35	0.14
vul_ing	0.08	0.00	0.04	0.09	0.05	0.00	0.05	0.09	0.00	0.07	0.11	0.05	0.04
no_pobv	0.38	0.00	0.06	0.49	0.02	0.00	0.03	0.47	0.00	0.32	0.15	0.55	0.01

Tabla 6b. **Estrato 2:** Algunos perfiles socioeconómicos de las cuatro Agrupaciones*.

	Agrupación 1*				Agrupación 2*				Agrupación 3*				Agrupación 4*		
	G1	G2	G3	G4	G1	G2	G3	G4	G1	G2	G3	G4	G1	G2	G3
ictSTpc	1637.70	2327.05	5347.86	2149.98	3035.46	5010.34	1990.89	1664.05	2256.10	5007.88	1373.44	2329.58	2487.21	5703.39	1289.69
ing	1305.14	1844.98	4064.04	1763.02	2146.82	4108.32	1327.78	1358.40	1707.45	3919.85	923.48	1907.15	1909.33	4469.81	929.83
pLB	0.77	0.65	0.31	0.67	0.58	0.30	0.75	0.76	0.68	0.33	0.86	0.65	0.60	0.25	0.86
pLBM	0.50	0.36	0.10	0.38	0.29	0.10	0.45	0.49	0.39	0.11	0.62	0.38	0.28	0.07	0.60
asaludp	0.53	0.50	0.28	0.49	0.61	0.18	0.54	0.52	0.50	0.29	0.58	0.54	0.49	0.23	0.59
rezedu15	0.11	0.05	0.03	0.04	0.05	0.03	0.06	0.10	0.03	0.01	0.05	0.99	0.05	0.02	0.09
rezedu16m	0.34	0.32	0.11	0.37	0.33	0.07	0.41	0.33	0.41	0.03	0.54	0.41	0.29	0.09	0.43
ali17	2.52	0.25	0.01	0.26	0.02	0.07	0.31	2.45	0.55	0.07	0.89	0.74	0.30	0.09	0.87
ali18m	2.66	0.53	0.00	2.45	0.00	0.38	1.71	2.40	1.22	0.17	1.57	1.05	0.80	0.30	1.45
segsoc	0.76	0.70	0.39	0.68	0.81	0.27	0.76	0.77	0.71	0.41	0.89	0.72	0.65	0.34	0.87
cv	0.08	0.04	0.01	0.05	0.03	0.01	0.05	0.08	0.04	0.00	0.54	0.05	0.02	0.01	0.08
sbv	0.10	0.05	0.01	0.08	0.04	0.01	0.08	0.09	0.06	0.01	0.27	0.08	0.04	0.01	0.11
ic_rezedu	0.61	0.55	0.24	0.59	0.55	0.21	0.65	0.59	0.66	0.14	0.77	1.00	0.51	0.21	0.70
ic_asalud	0.45	0.41	0.30	0.40	0.50	0.22	0.43	0.43	0.41	0.31	0.45	0.47	0.42	0.27	0.44
ic_segsoc	0.87	0.82	0.55	0.81	0.88	0.47	0.87	0.87	0.82	0.57	0.96	0.83	0.79	0.51	0.94
ic_cv	0.28	0.12	0.02	0.14	0.09	0.03	0.17	0.26	0.16	0.00	1.00	0.15	0.08	0.02	0.26
ic_sbv	0.30	0.18	0.04	0.25	0.14	0.05	0.26	0.28	0.20	0.05	0.68	0.25	0.13	0.04	0.34
ic_ali	1.00	0.07	0.00	0.92	0.00	0.08	0.48	1.00	0.41	0.02	0.51	0.35	0.24	0.07	0.51
pobreza_e	0.26	0.07	0.00	0.17	0.05	0.00	0.15	0.24	0.11	0.00	0.39	0.23	0.04	0.00	0.25
pobreza_m	0.52	0.53	0.13	0.51	0.47	0.12	0.60	0.53	0.54	0.15	0.46	0.42	0.48	0.09	0.63
vul_car	0.22	0.24	0.35	0.30	0.43	0.27	0.20	0.23	0.29	0.31	0.15	0.21	0.32	0.37	0.10
vul_ing	0.00	0.08	0.10	0.01	0.02	0.14	0.03	0.00	0.03	0.12	0.00	0.04	0.09	0.08	0.02
no_pobv	0.00	0.09	0.42	0.01	0.03	0.47	0.02	0.00	0.03	0.42	0.00	0.11	0.06	0.46	0.00

Tabla 6c. **Estrato 3:** Algunos perfiles socioeconómicos de las cuatro Agrupaciones*.

	Agrupación 1*					Agrupación 2*			Agrupación 3*				Agrupación 4*		
	G1	G2	G3	G4	G5	G1	G2	G3	G1	G2	G3	G4	G1	G2	G3
ictSTpc	1644.07	3900.75	1874.62	1311.14	1351.42	1739.38	4609.74	1613.49	1940.32	1162.62	3135.90	1438.46	4826.12	1058.42	2289.04
ing	1133.28	3215.62	1590.65	1070.55	1059.52	1267.13	3877.34	1329.80	1661.85	836.09	2527.78	1145.00	4121.03	783.75	1719.08
pLB	0.81	0.42	0.72	0.83	0.83	0.77	0.31	0.77	0.70	0.89	0.54	0.81	0.28	0.90	0.65
pLBM	0.52	0.16	0.45	0.58	0.59	0.48	0.10	0.52	0.43	0.65	0.27	0.57	0.08	0.67	0.32
asaludp	0.64	0.31	0.53	0.54	0.55	0.56	0.28	0.53	0.49	0.59	0.42	0.57	0.26	0.60	0.50
rezedu15	0.06	0.03	0.03	0.08	0.12	0.06	0.03	0.06	0.74	0.01	0.01	0.08	0.03	0.07	0.04
rezedu16m	0.39	0.18	0.50	0.37	0.39	0.38	0.10	0.42	0.33	0.52	0.22	0.47	0.09	0.50	0.29
ali17	0.00	0.09	0.00	1.41	3.00	0.26	0.09	1.17	0.71	0.82	0.22	0.81	0.13	0.76	0.32
ali18m	0.27	0.22	2.48	1.61	2.80	0.39	0.19	2.50	0.90	1.63	0.50	1.32	0.38	1.25	0.73
segsoc	0.95	0.41	0.70	0.84	0.84	0.79	0.42	0.76	0.77	0.86	0.60	0.83	0.40	0.88	0.69
cv	0.05	0.02	0.06	0.09	0.13	0.05	0.01	0.09	0.00	0.00	0.00	0.32	0.01	0.09	0.03
sbv	0.12	0.03	0.16	0.14	0.17	0.11	0.02	0.15	0.12	0.23	0.03	0.22	0.02	0.18	0.07
ic_rezedu	0.65	0.34	0.67	0.64	0.66	0.63	0.24	0.65	1.00	0.75	0.39	0.74	0.23	0.78	0.49
ic_asalud	0.46	0.30	0.38	0.41	0.45	0.41	0.30	0.41	0.39	0.40	0.36	0.45	0.29	0.43	0.39
ic_segsoc	1.00	0.59	0.80	0.92	0.93	0.88	0.60	0.86	0.86	0.92	0.73	0.91	0.59	0.94	0.80
ic_cv	0.16	0.06	0.21	0.27	0.36	0.16	0.04	0.27	0.00	0.00	0.00	1.00	0.04	0.28	0.11
ic_sbv	0.36	0.12	0.41	0.40	0.47	0.33	0.08	0.41	0.33	0.69	0.10	0.57	0.07	0.50	0.22
ic_ali	0.00	0.02	0.87	0.69	1.00	0.08	0.02	0.95	0.33	0.59	0.13	0.47	0.10	0.43	0.22
pobreza_e	0.15	0.01	0.27	0.27	0.35	0.13	0.00	0.31	0.14	0.32	0.02	0.40	0.00	0.32	0.05
pobreza_m	0.66	0.24	0.45	0.59	0.51	0.60	0.15	0.48	0.57	0.57	0.41	0.42	0.14	0.59	0.52
vul_car	0.18	0.36	0.26	0.12	0.14	0.20	0.38	0.20	0.19	0.11	0.31	0.18	0.41	0.08	0.30
vul_ing	0.01	0.12	0.01	0.01	0.00	0.05	0.11	0.00	0.04	0.00	0.08	0.00	0.09	0.01	0.08
no_pobv	0.00	0.27	0.02	0.00	0.00	0.02	0.36	0.01	0.05	0.00	0.17	0.00	0.36	0.00	0.05

Tabla 6d. **Estrato 4:** Algunos perfiles socioeconómicos de las cuatro Agrupaciones*.

	Agrupación 1*			Agrupación 2*			Agrupación 3*					Agrupación 4*			
	G1	G2	G3	G1	G2	G3	G1	G2	G3	G4	G5	G1	G2	G3	G4
ictSTpc	3106.84	1146.04	1144.30	1251.58	881.11	4083.02	769.10	1375.68	3562.80	673.01	532.40	603.86	4661.28	1249.96	2091.62
ing	2269.36	826.28	689.58	819.12	587.52	3036.74	474.25	869.47	2694.25	443.46	379.97	373.53	3567.61	774.63	1432.38
pLB	0.42	0.79	0.82	0.78	0.86	0.26	0.91	0.75	0.33	0.92	0.94	0.95	0.20	0.79	0.53
pLBM	0.21	0.60	0.61	0.56	0.68	0.09	0.74	0.51	0.14	0.77	0.81	0.82	0.06	0.52	0.25
asaludp	0.43	0.58	0.58	0.57	0.58	0.36	0.60	0.60	0.34	0.66	0.63	0.60	0.31	0.58	0.52
rezedu15	0.04	0.07	0.06	0.05	0.11	0.04	0.05	0.05	0.05	0.94	0.10	0.08	0.04	0.05	0.05
rezedu16m	0.31	0.55	0.52	0.52	0.50	0.23	0.60	0.48	0.28	0.69	0.70	0.63	0.19	0.50	0.38
ali17	0.24	1.30	0.15	0.31	2.55	0.15	1.12	0.19	0.46	1.03	0.95	0.82	0.25	0.45	0.34
ali18m	0.53	2.41	0.26	0.89	2.62	0.29	1.49	0.76	0.72	1.25	1.50	1.27	0.49	1.00	0.74
segsoc	0.63	0.85	0.86	0.83	0.93	0.52	0.91	0.90	0.44	0.98	0.94	0.92	0.48	0.84	0.72
cv	0.03	0.09	0.07	0.07	0.11	0.02	0.08	0.05	0.02	0.08	0.55	0.11	0.02	0.05	0.04
sbv	0.12	0.29	0.26	0.25	0.29	0.08	0.46	0.15	0.10	0.60	0.64	0.37	0.07	0.22	0.13
ic_rezedu	0.53	0.76	0.76	0.75	0.74	0.43	0.83	0.72	0.48	1.00	0.89	0.88	0.38	0.73	0.59
ic_asalud	0.33	0.40	0.39	0.38	0.41	0.31	0.38	0.41	0.29	0.50	0.45	0.40	0.30	0.37	0.39
ic_segsoc	0.75	0.91	0.92	0.90	0.97	0.66	0.96	0.96	0.58	1.00	0.98	0.97	0.64	0.90	0.81
ic_cv	0.09	0.28	0.22	0.21	0.33	0.06	0.30	0.16	0.08	0.34	1.00	0.35	0.05	0.17	0.12
ic_sbv	0.34	0.70	0.64	0.63	0.72	0.24	0.91	0.51	0.29	1.00	1.00	0.84	0.21	0.59	0.40
ic_ali	0.06	0.97	0.02	0.27	1.00	0.04	0.56	0.19	0.23	0.42	0.55	0.44	0.13	0.31	0.21
pobreza_e	0.03	0.34	0.20	0.21	0.41	0.01	0.41	0.14	0.02	0.56	0.68	0.45	0.01	0.14	0.04
pobreza_m	0.26	0.36	0.49	0.45	0.38	0.13	0.42	0.49	0.17	0.26	0.23	0.44	0.09	0.51	0.32
vul_car	0.52	0.30	0.28	0.32	0.21	0.58	0.16	0.36	0.57	0.17	0.08	0.10	0.59	0.34	0.57
vul_ing	0.02	0.00	0.01	0.01	0.00	0.03	0.00	0.01	0.04	0.00	0.00	0.00	0.02	0.01	0.03
no_pobv	0.16	0.00	0.01	0.01	0.00	0.26	0.00	0.00	0.21	0.00	0.00	0.00	0.29	0.00	0.05

La Agrupación 3*, en contraste a las otras tres agrupaciones, forma grupos diferenciados por carencias específicas de pobreza multidimensional, es decir, los grupos no están ordenados de manera global por todas las carencias. Por ejemplo, en el Estrato 4 (Tabla 6d), el grupo 3 es el menos pobre de manera global, excepto en la carencia de alimentación en menores (ali17) en donde ocupa el segundo lugar de pobreza. En cambio, el grupo 4 es el segundo más pobre de acuerdo con el ingreso (estimado y observado), es el más pobre en acceso a servicios de salud (asaludp), el más pobre en rezago educativo en menores (rezedu15) y el tercero más pobre en carencia alimentaria en adultos (ali18m).

8. SELECCIÓN DE LA AGRUPACIÓN

A lo largo de este documento se han presentado cuatro propuestas de agrupaciones. Estas agrupaciones usan técnicas y criterios distintos, lo que da lugar a agrupaciones con características distintas también. Las cuatro propuestas tienen ventajas y desventajas que se resumen a continuación.

Agrupación 1*

Forma los grupos usando principalmente las variables de carencias de alimentación en menores (ali17) y en adultos (ali18m). Forma 15 grupos en total considerando todos los estratos de tamaño de localidad. El grupo más pequeño tiene 1 millón de personas y el más grande 44 millones. Los grupos son relativamente fáciles de identificar debido a que sus árboles tienen en total, en todos los estratos, 25 nodos terminales. Los grupos formados están ordenados, dentro de cada estrato, en cuanto a la pobreza multidimensional global, dominada por ingreso.

Agrupación 2*

Forma los grupos usando principalmente las probabilidades de que el ingreso observado esté por debajo de las líneas de bienestar (pLB) y bienestar mínimo (pLBM) junto con la variable de carencia de alimentación en adultos (ali18m). Tiene 13 grupos en total en todos los estratos. El tamaño de los grupos varía entre el grupo más pequeño de 2 millones de personas y el más grande de 33 millones. Los grupos son relativamente fáciles de identificar debido a que sus árboles de clasificación tienen en total

24 nodos terminales. Los grupos formados están ordenados dentro de cada estrato en la mayoría de las variables, excepto en acceso a servicios de salud (asaludp) y en carencia por acceso a la seguridad social (segsoc).

Agrupación 3*

Forma sus grupos usando todas las variables de pobreza multidimensional. Tiene 17 grupos en total y el tamaño de sus grupos varía entre el más pequeño de 250 mil personas y el más grande de 33 millones. Los grupos son claramente identificables e incluyen muchos criterios. Los árboles tienen en total 31 nodos terminales. Los grupos formados no están ordenados por sus condiciones de pobreza global sino que están diferenciados por sus condiciones específicas de pobreza.

Agrupación 4*

Forma los grupos usando exclusivamente las variables de ingreso (ing, pLB y pLBM). Tiene 13 grupos en total con tamaños que varían entre los 4 millones y los 24 millones de personas. Los grupos son muy fáciles de identificar debido a que sus árboles tienen 13 nodos terminales en total. Los grupos formados se encuentran ordenados, dentro de cada estrato, por las condiciones de pobreza global dominada por ingreso.

Agrupación Seleccionada

Tomando en cuenta todas estas características tanto estadísticas como sustantivas, además de los perfiles presentados en la sección anterior, se considera que la **Agrupación 3*** es la que permite diferenciar mejor a los hogares en cuanto a los distintos indicadores de la pobreza. En consecuencia, al usar esta agrupación para focalizar a los programas se tiene una mejor identificación de las necesidades específicas de la población y es posible asignar de una mejor manera las personas y los hogares a los distintos programas que ofrece la SEDESOL. En consecuencia, sugiere utilizar a la Agrupación 3* como Modelo Único de Clasificación (MUC).

9. CREACIÓN DE INDICADORES

Para facilitar la intervención operativa de los programas, es posible crear un indicador que permita ordenar a los hogares con mayor propensión a recibir programas dentro de cada uno de los 17 grupos formados con la Agrupación 3*. Adicionalmente, si es necesario comparar a los hogares de un grupo contra otro pero dentro del mismo estrato, se propone un segundo indicador que permita hacerlo. Finalmente, se propone un indicador poblacional que permite comparar a los hogares de distintos grupos y de distintos estratos.

Estos indicadores resumen se construyen con la técnica estadística llamada Análisis de Componentes Principales. Para la construcción de los indicadores primero se estandarizan las 11 variables de intensidades de carencias considerando a todos los hogares del país. A cada variable se le resta su media y se divide esta diferencia entre su desviación estándar.⁶ Las medias y las varianzas muestrales de las 11 variables se muestran en la Tabla 7.

Tabla 7. Valores para estandarización. Media y desviación estándar.

Variable	Media	Desv.Est.
pLB	0.5177	0.2837
pLBM	0.2882	0.2548
asaludp	0.4030	0.3229
rezedu15	0.0440	0.1786
rezedu16m	0.2626	0.3616
ali17	0.3440	0.7754
ali18m	0.7255	1.0238
segsoc	0.5802	0.4380
cv	0.0342	0.1071
sbv	0.0761	0.1662
ling	7.5827	0.8588

⁶Si la variable original se representa por x_i , su media y desviación estándar estimadas es \bar{x} y s , respectivamente, entonces la variable estandarizada de x_i es $z_i = (x_i - \bar{x})/s$.

Indicador intra-grupo.

Para crear este indicador que permita diferenciar a los hogares dentro de cada grupo, se utiliza exclusivamente la información de dichos hogares. En total se tienen 17 indicadores distintos, uno por cada grupo. Cada indicador es un promedio ponderado de las variables de intensidades de carencias estandarizadas. Los coeficientes de cada variable estandarizada para formar el indicador de cada grupo se pueden observar en las Tablas 8a y 8b.

De las tablas se nota que los coeficientes reportados varían de un grupo a otro. Como las variables son previamente estandarizadas a nivel nacional,⁷ los valores de los coeficientes indican la importancia de cada variable en la determinación del indicador. Por ejemplo, en el grupo1 del Estrato 1 la variable más importante es el rezago educativo en menores y la menos importante es la carencia por calidad y espacios de la vivienda (cv), en cambio en el grupo 2 del mismo Estrato 1, la variable más importante es el acceso a servicios de salud y la seguridad social y las menos importantes son acceso a servicios básicos de vivienda y carencia alimentaria en adultos. Estas variables con mayor peso en cada indicador son las que discriminan internamente dentro de cada grupo.

Tabla 8a. Coeficientes para formar los indicadores intra-grupo en los Estratos 1 y 2.

Variable	E1 G1	E1 G2	E1 G3	E1 G4	E2 G1	E2 G2	E2 G3	E2 G4
pLB*	-0.1896	0.0113	-0.0075	-0.0943	-0.2577	-0.2468	-0.0603	-0.3039
pLBM*	-0.2211	0.0194	-0.0139	-0.0390	-0.3124	-0.1352	-0.1261	-0.3465
asaludp*	0.0091	-0.7285	-0.0711	-0.7117	-0.1751	-0.6874	-0.1307	-0.1926
rezedu15*	-0.8720	0.0105	-0.0875	-0.0136	-0.0444	0.0052	-0.0091	0.0420
rezedu16m*	-0.1831	0.0370	-0.1039	-0.0934	0.1525	-0.0189	-0.1463	-0.3276
ali17*	-0.0809	0.0114	0.6222	0.0160	-0.6251	-0.0098	-0.5122	-0.4138
ali18m*	-0.2972	0.0046	0.3353	0.0059	-0.5038	-0.0224	-0.4494	-0.3462
segsoc*	-0.0323	-0.6834	-0.0626	-0.6824	-0.2357	-0.6356	-0.0268	-0.2426

⁷ Para diferenciar las variables estandarizadas, éstas se denotan con (*).

cv*	0.0000	-0.0085	-0.6802	-0.0127	-0.1021	0.0000	-0.2696	-0.4220
sbv*	-0.0325	0.0001	-0.1003	-0.0031	-0.1242	-0.0181	-0.6294	-0.1851
ling*	0.1559	0.0068	0.0100	0.0896	0.2330	0.2072	0.0896	0.2811

Tabla 8b. Coeficientes para formar los indicadores intra-grupo en los Estratos 3 y 4.

Variable	E3 G1	E3 G2	E3 G3	E3 G4	E4 G1	E4 G2	E4 G3	E4 G4	E4 G5
pLB*	-0.3644	0.0020	-0.4514	-0.2530	-0.0036	-0.3084	-0.2573	-0.0633	-0.0196
pLBM*	-0.4273	0.0022	-0.3910	-0.3855	-0.0109	-0.4578	-0.1777	-0.1545	-0.0610
asaludp*	-0.2238	0.0274	-0.4733	-0.1398	0.0226	-0.0124	-0.2670	-0.0464	-0.0458
rezedu15*	0.0530	-0.0051	-0.0090	-0.1292	-0.0825	-0.0583	-0.1693	0.1087	-0.4171
rezedu16m*	-0.2953	0.2150	-0.2295	-0.2074	0.1347	-0.5282	-0.1536	-0.1873	0.0087
ali17*	-0.4264	-0.7376	-0.0677	-0.4830	-0.7346	-0.0482	-0.5991	-0.7771	-0.7708
ali18m*	-0.3405	-0.4812	-0.0709	-0.3699	-0.3875	-0.2380	-0.4590	-0.5352	-0.4213
segsoc*	-0.2816	-0.0263	-0.4675	-0.1550	-0.0527	0.0873	-0.2678	-0.0010	-0.0182
cv*	0.0000	0.0000	0.0000	-0.2693	-0.0319	-0.3945	-0.1964	-0.0985	-0.1803
sbv*	-0.2302	0.4204	-0.0202	-0.3937	0.5298	-0.2896	-0.2204	-0.0368	-0.0918
ling*	0.3383	0.0009	0.3716	0.2940	0.0129	0.3265	0.2241	0.1476	0.0830

Indicadores intra-estrato y nacional.

Para crear los indicadores que permitan diferenciar a los hogares de distintos grupos pero dentro de un mismo estrato, se usa la información de todos los hogares en el mismo estrato. En la Tabla 9 se reportan los coeficientes para formar los indicadores ponderados de las variables estandarizadas para cada uno de los cuatro estratos. Al igual que en los indicadores intra-grupo es posible comparar la importancia que tienen las variables en la definición del indicador. Por ejemplo, en el Estrato 1 las variables más importantes son la probabilidad de estar por debajo de la línea de bienestar y los accesos a servicios de salud y la menos importante es el acceso a los servicios básicos de vivienda. Por otro lado, en el Estrato 4 la variable más importante es acceso a servicios básicos de vivienda y la probabilidad de estar por debajo de la línea de bienestar mínimo y la menos importante es el rezago educativo en menores.

Tabla 9. Coeficientes para formar los indicadores intra-estrato e indicador nacional.

Variable	E1	E2	E3	E4	Nacional
pLB*	-0.4386	-0.4164	-0.4092	-0.3770	-0.4232
pLBM*	-0.3330	-0.3901	-0.4454	-0.4519	-0.4262
asaludp*	-0.4268	-0.3265	-0.2688	-0.1496	-0.2789
rezedu15*	-0.1197	-0.0993	-0.0745	-0.0782	-0.0789
rezedu16m*	-0.1912	-0.2391	-0.2629	-0.2337	-0.2570
ali17*	-0.2320	-0.2845	-0.2565	-0.2179	-0.2133
ali18m*	-0.3028	-0.3268	-0.2754	-0.2144	-0.2440
segsoc*	-0.3926	-0.3396	-0.2889	-0.1735	-0.2911
cv*	-0.1365	-0.2266	-0.2427	-0.2910	-0.1981
sbv*	-0.0491	-0.1433	-0.2523	-0.4655	-0.2757
ling*	0.3782	0.3546	0.3736	0.3818	0.4292

En la última columna de la Tabla 9 se incluyen también los pesos para formar un indicador nacional que permita comparar a todos los hogares del país por sus condiciones generales de pobreza multidimensional medida en las dimensiones definidas en este estudio.

Es importante mencionar que estos indicadores intra-grupo, intra-estrato o nacional permiten establecer un orden en la necesidad de atención de los hogares en cuanto a sus características de pobreza multidimensional, esto es un indicador de la necesidad de atención; sin embargo, para algunos programas específicos es posible usar variables operativas de interés para complementar la focalización y priorizar la intervención de una mejor manera.

10. VENTAJAS Y DESVENTAJAS DEL MUC

Una de las ventajas mayores del MUC consiste en que el modelo está pensado para focalizar con mayor precisión a los hogares que potencialmente requieren de atención prioritaria para su posible asignación a uno o más programas sociales. De esta manera el modelo puede identificar a los hogares que potencialmente pueden ser receptores de algún programa social. Además con la creación de indicadores se genera un puntaje que mide la propensión de un hogar para ser receptor de uno o

varios programas, dependiendo de los criterios establecidos por el techo presupuestal.

Previamente al desarrollo del MUC se conocía con alto grado de incertidumbre si un hogar es elegible para la asignación de un programa social y posiblemente un hogar podría recibir uno o varios programas sociales, pero se carecía de un modelo que pudiera hacer tal distinción. El MUC está diseñado para automatizar el proceso de selección, con la posibilidad de minimizar sesgos o errores en la elegibilidad de hogares a programas sociales.

Para este efecto el MUC utiliza la muestra ENIGH como fuente de información, así como los criterios de pobreza multidimensional establecidos por el CONEVAL, por tanto no sólo se basa en el ingreso sino en el enfoque multidimensional de la pobreza.