

SECRETARÍA DISTRITAL DE PLANEACIÓN

ANALIZAR LA ESTRATIFICACIÓN COMO INSTRUMENTO DE FOCALIZACIÓN  
JUNTO CON OTROS INSTRUMENTOS DE POLÍTICA PÚBLICA PARA  
DETERMINAR SU EFICIENCIA

(Contrato No. 707 de 2023)

ANÁLISIS DESCRIPTIVO DEL ÍNDICE PROPUESTO  
(Producto 4, Vr. 7)

Bogotá, 2 de abril de 2024

## Contenido

Contenido .....	2
INTRODUCCIÓN .....	10
1. DESCRIPCIÓN DE LAS VARIABLES INVOLUCRADAS EN EL ÍNDICE .....	12
1.1. LOS NIVELES DE INGRESO EN EL DESARROLLO PROGRESIVO DEL RUI 12	
1.2. LOS COMPONENTES DEL ÍNDICE .....	15
1.2.1. Los perceptores por tipo de ingreso .....	16
1.2.2. Diferencias en los ingresos por grupos poblacionales y en el aporte de los distintos tipos de ingresos .....	19
2. RESULTADO CUANTITATIVO DEL ÍNDICE PROPUESTO, CON EL ANÁLISIS DE LA SEMAFORIZACIÓN FRENTE A LA ASIGNACIÓN DE SUBSIDIOS PARA LA PRESTACIÓN DE SERVICIOS PÚBLICOS DOMICILIARIOS.....	32
2.1 Resultado cuantitativo del índice propuesto. ....	32
2.2 Semaforización.....	36
3. CARACTERIZACIÓN DE LOS GRUPOS DEL ÍNDICE DE INGRESOS.....	38
3.1. Características demográficas .....	38
3.2. Condiciones de vulnerabilidad.....	41
3.3. Índice de pobreza multidimensional (IPM) y sus componentes por grupos poblacionales de los modelos 1 a 5 del índice de ingresos. ....	48
3.4. Distribución por grupos de los modelos 1 a 5 del índice RUI de ingresos por UPZ .....	52
3.5 UPZ por estrato vs clasificación por índice RUI de ingresos Modelo 2 .....	68
4. RELACIÓN CUANTITATIVA DEL ÍNDICE CONSTRUIDO CON LA ESTRATIFICACIÓN VIGENTE .....	73
5. ESTIMACIÓN DE LOS ERRORES DE INCLUSIÓN Y DE EXCLUSIÓN.....	75
6. LOS REGISTROS ADMINISTRATIVOS (RA) DISPONIBLES Y EL REGISTRO UNIVERSAL DE INGRESOS (RUI).....	77
6.1. Las diferencias entre la información de la PILA y la de la EMB 2021 sobre empleo y aportes a la seguridad social en salud .....	77
6.2. Caracterización externa de los grupos de trabajadores identificados en la PILA.....	81
6.3. La información sobre ingresos del SISBÉN IV.....	84
6.3.1. La información sobre ingresos .....	84
6.3.2. Estimación de los ingresos de las unidades de gasto a partir de variables socioeconómicas del SISBEN IV utilizando la técnica de regresión .....	85

7. RECOMENDACIONES DE POLÍTICA PÚBLICA.....	93
7.1 Recomendaciones de política sobre Habeas Data .....	96
8. CONCLUSIONES Y RECOMENDACIONES .....	101
REFERENCIAS.....	104
APÉNDICE 1. CONSTRUCCIÓN DEL ÍNDICE CON RUTINAS .....	106
APÉNDICE 2. MODELO DE ESTIMACIÓN DE INGRESOS.....	126

## ÍNDICE DE CUADROS

Cuadro 1. Semaforización por Grupos de clasificación según IRUI y Nivel de ingresos para el Modelo propuesto 1 .....	37
Cuadro 2. Porcentaje de personas por número de situaciones de vulnerabilidad que afectan sus hogares por grupos de ingreso del modelo 2- Bogotá- 2021. ....	48
Cuadro 3. Indicadores de los componentes del Índice de Pobreza Multidimensional-Modelo 1 del Índice de Ingresos- Bogotá- 2021 .....	50
Cuadro 4. Indicadores de los componentes del Índice de Pobreza Multidimensional-Modelo 2 del Índice de Ingresos- Bogotá- 2021 .....	50
Cuadro 5. Indicadores de los componentes del Índice de Pobreza Multidimensional-Modelo 3 del Índice de Ingresos- Bogotá- 2021 .....	51
Cuadro 6. Indicadores de los componentes del Índice de Pobreza Multidimensional-Modelo 4 del Índice de Ingresos- Bogotá- 2021 .....	51
Cuadro 7. Indicadores de los componentes del Índice de Pobreza Multidimensional-Modelo 5 del Índice de Ingresos- Bogotá- 2021 .....	51
Cuadro 8. UPZ según estrato y clasificación por niveles de ingreso Modelo 2.....	68
Cuadro 9. UPZ según estrato y clasificación por niveles de ingreso Modelo 2.....	69
Cuadro 10. Población por ubicación en los estratos socioeconómicos actuales y los grupos del modelo 2 del ingreso secundario del RUI- 2021.....	73
Cuadro 11. Errores de inclusión y de exclusión de la estratificación vigente y los grupos de los modelos 1 a 5 – Bogotá – 2021.....	75
Cuadro 12. Información sobre número de cotizantes y trabajadores y remuneración laboral de la PILA y la EMB 2021. Bogotá-2021 .....	80
Cuadro 13. Indicadores de vulnerabilidad de los asalariados afiliados al régimen contributivo de salud. Porcentajes de personas pertenecientes a hogares con la característica del indicador dentro de cada grupo de ingresos. Bogotá-2021 .....	82
Cuadro 14. Indicadores de vulnerabilidad de los trabajadores independientes afiliados al régimen contributivo de salud. Porcentajes de personas pertenecientes a hogares con la característica del indicador dentro de cada grupo de ingresos. Bogotá-2021 .....	82
Cuadro 15. Indicadores de vulnerabilidad de los trabajadores domésticos afiliados al régimen contributivo de salud. pertenecientes a hogares con la característica del indicador dentro de cada grupo de ingresos. Bogotá-2021 .....	83
Cuadro 16. Indicadores de vulnerabilidad de los afiliados a los regímenes especiales y de excepción de salud. pertenecientes a hogares con la característica del indicador dentro de cada grupo de ingresos. Bogotá-2021 .....	83
Cuadro 17. Porcentaje de personas registradas en el SISBÉN IV, sin ingresos, según concepto de ingreso. Bogotá. ....	84
Cuadro 18. Variables utilizadas en el modelo de regresión utilizado para estimar el ingreso per cápita de las unidades de gasto con la encuesta del SISBÉN IV. ....	86
Cuadro 19. Coeficientes obtenidos del modelo de regresión aplicado a la EMB 2021 para estimar el ingreso per cápita de las unidades de gasto con información del SISBÉN IV- 2021 .....	87
Cuadro 20. Distribución de la población registrada en el SISBÉN IV, por grupos de niveles de ingreso per cápita del modelo 2- 2021. ....	88
Cuadro 21. Programas del orden distrital que involucran estratificación social .....	96

## ÍNDICE DE GRÁFICOS

Gráfico 1. Perceptores de ingresos de 12 y más años por tipo de ingreso- Bogotá-2021	17
Gráfico 2. Promedio mensual de los tipos de ingresos de personas de 12 y más años por tipo de ingreso- Bogotá-2021 (en pesos corrientes).....	17
Gráfico 3. Porcentaje de perceptores de ingresos, por sexo y grupos de edad- Bogotá-2021 .....	18
Gráfico 4. Porcentaje de perceptores de ingresos de 12 y más años por Localidad-Bogotá-2021 .....	19
Gráfico 5. Participación de los distintos tipos de ingreso en el ingreso secundario total personal- Bogotá-2021 .....	20
Gráfico 6. Ingreso secundario total, personal, por sexo-Bogotá-2021 .....	21
Gráfico 7. Participación de los distintos tipos de ingreso en el ingreso secundario total personal, por sexo- Bogotá-2021 .....	21
Gráfico 8. Ingreso secundario total, personal, por grupos de edad-Bogotá-2021 (pesos corrientes).....	22
Gráfico 9. Participación porcentual de los distintos tipos de ingreso en el ingreso secundario total personal, por grupos de edad- Bogotá-2021 .....	23
Gráfico 10. Participación de los distintos tipos de ingreso en el ingreso secundario total personal, por zonas urbana (cabecera) y rural (resto)- Bogotá-2021 .....	24
Gráfico 11. Participación de los distintos tipos de ingreso personal en el ingreso total- Bogotá-2021 .....	25
Gráfico 12. Ingreso secundario total, personal, por Localidad-Bogotá-2021 .....	25
Gráfico 13. Participación de los distintos tipos de ingreso personal en el ingreso total, por Localidad- Bogotá-2021 .....	26
Gráfico 14. Ingreso per cápita secundario , por grupos del modelo 2 del índice de ingresos- Bogotá-2021 (pesos corrientes).....	27
Gráfico 15. Participación porcentual de los distintos tipos de ingreso en el ingreso secundario total personal, por grupos del modelo 2 del índice de ingresos- Bogotá-2021	28
Gráfico 16. Ingreso secundario total, personal, por participación en el programa de Familias en Acción- -Bogotá-2021 (pesos corrientes) .....	29
Gráfico 17. Participación porcentual de los distintos tipos de ingreso en el ingreso secundario total personal, por participación en el programa de Familias en Acción- Bogotá-2021 .....	30
Gráfico 18. Ingreso secundario total, personal, por participación en el programa de Adultos Mayores-Bogotá-2021 (pesos corrientes) .....	30
Gráfico 19. Participación porcentual de los distintos tipos de ingreso en el ingreso secundario total personal, por participación en el programa de Adultos Mayores- Bogotá-2021 .....	31
Gráfico 20. Índice promedio RUI por sexo- Bogotá-2021 .....	32
Gráfico 21. Índice promedio RUI por edades simples- Bogotá-2021 .....	33
Gráfico 22. Índice RUI promedio por grupos de edad- Bogotá-2021 .....	33

Gráfico 23. Índice promedio RUI por estratos socioeconómicos vigentes-Bogotá- 2021..	34
Gráfico 24. Índice promedio RUI por zonas-Bogotá- 2021.....	35
Gráfico 25. Índice RUI promedio por localidades- Bogotá-2021.....	36
Gráfico 26. Razón de masculinidad, por grupos del modelo 2 del índice RUI -Bogotá-2021 .....	39
Gráfico 27. Edad promedio, por grupos del modelo 2 del índice RUI -Bogotá-2021.....	39
Gráfico 28. Promedio del número de personas por hogar, por grupos del modelo 2 del índice RUI -Bogotá-2021.....	40
Gráfico 29. Porcentaje de población indígena y negra, por grupos del modelo 2 del índice RUI -Bogotá-2021.....	41
Gráfico 30. Porcentaje de personas en hogares con menores de edad por pertenencia a los grupos del índice de ingreso del Modelo 2- Bogotá - 2021 .....	42
Gráfico 31. Porcentaje de personas en hogares con mayores de 60 años sin ingreso por pertenencia a los grupos del índice de ingreso del Modelo 2- Bogotá - 2021.....	43
Gráfico 32. Porcentaje de población con discapacidad por grupos del modelo 2 del índice RUI -Bogotá-2021.....	44
Gráfico 33. Porcentaje de personas en hogares con discapacidad por pertenencia a los grupos del índice de ingreso del Modelo 2- Bogotá – 2021.....	44
Gráfico 34. Porcentaje de población sin afiliación a la Seguridad Social en Salud, por grupos del modelo 2 del índice RUI -Bogotá-2021.....	45
Gráfico 35. Porcentaje de población en viviendas con déficit habitacional, por grupos del modelo 2 del índice RUI -Bogotá-2021 .....	46
Gráfico 36. Porcentaje de personas en hogares con hacinamiento por pertenencia a los grupos del índice de ingreso del Modelo 2- Bogotá - 2021 .....	46
Gráfico 37. Porcentaje de personas en hogares con hacinamiento por pertenencia a los grupos del índice de ingreso del Modelo 2- Bogotá - 2021 .....	47
Gráfico 38. UPZ 103 Ingresos promedios por estrato (tarifa de energía) vs clasificación Modelo 2.....	70
Gráfico 39. UPZ 108 y 111 Ingresos promedios por estrato (tarifa de energía) vs clasificación Modelo 2.....	70
Gráfico 40. UPZ 14 Ingresos promedios por estrato (tarifa de energía) vs clasificación Modelo 2.....	71
Gráfico 41. UPZ 85 Ingresos promedios por estrato (tarifa de energía) vs clasificación Modelo 2.....	72
Gráfico 42. Distribución de la población de Bogotá por grupos del Modelo 2 y la estratificación vigente- 2021 .....	74
Gráfico 43. Distribución porcentual de la población de Bogotá por grupos del Modelo 2 y la estratificación vigente- 2021 .....	74
Gráfico 44. Distribución porcentual de los hogares del SISBÉN IV, por grupos de ingreso RUI del modelo 2 del RUI-2021 .....	88
Gráfico 45. Población en los grupos 1, 2 y 3 del modelo 2 del Índice RUI, con la EMB 2021 y con la estimación a partir del SISBÉN IV .....	90
Gráfico 46. Ingreso promedio por unidad de gasto de los grupos 1, 2 y 3 del modelo 2 del índice RUI según la EMP 2021 y el SISBÉN IV .....	90
Gráfico 47. Promedios de personas por hogar por grupos del modelo 2, obtenidos de la EMP 2021 y del SISBÉN IV .....	91

<i>Gráfico 48. Edad promedio por grupos del modelo 2, obtenidas de la EMP 2021 y del SISBÉN IV .....</i>	<i>91</i>
<i>Gráfico 49. Razón de masculinidad, por grupos del modelo 2, obtenidas de la EMP 2021 y del SISBÉN IV .....</i>	<i>92</i>

## ÍNDICE DE ILUSTRACIONES

<i>Ilustración 1. Mapa de distribución de grupos del índice de ingreso por UPZ. Modelo 1 ..</i>	<i>54</i>
<i>Ilustración 2. Mapa de distribución de grupos del índice de ingreso por UPZ. Modelo 2 ..</i>	<i>55</i>
<i>Ilustración 3. Mapa de distribución de grupos del índice de ingreso por UPZ. Modelo 3 ..</i>	<i>56</i>
<i>Ilustración 4. Mapa de distribución de grupos del índice de ingreso por UPZ. Modelo 4 ..</i>	<i>57</i>
<i>Ilustración 5. Mapa de distribución de grupos del índice de ingreso por UPZ. Modelo 5 ..</i>	<i>58</i>
<i>Ilustración 6. Estratificación vigente por manzanas agrupadas por UPZ. ....</i>	<i>59</i>
<i>Ilustración 7. UPZ por grupo predominante del modelo 1 obtenido por interpolación IDW vs. Estratificación vigente .....</i>	<i>61</i>
<i>Ilustración 8. UPZ por grupo predominante del modelo 2 obtenido por interpolación IDW vs. Estratificación vigente .....</i>	<i>62</i>
<i>Ilustración 9. UPZ por grupo predominante del modelo 3 obtenido por interpolación IDW vs. Estratificación vigente .....</i>	<i>63</i>
<i>Ilustración 10. UPZ por grupo predominante del modelo 4 obtenido por interpolación IDW vs. Estratificación vigente .....</i>	<i>64</i>
<i>Ilustración 11. UPZ por grupo predominante del modelo 5 obtenido por interpolación IDW vs. Estratificación vigente .....</i>	<i>65</i>
<i>Ilustración 12. Valor promedio del índice de ingresos por UPZ .....</i>	<i>66</i>
<i>Ilustración 13. UPZ por grupo predominante del modelo con ingresos primarios obtenido por interpolación IDW vs. Estratificación .....</i>	<i>67</i>



## GLOSARIO DE TÉRMINOS

ADRES:	Administradora de los Recursos del Sistema General de Seguridad Social en Salud
CEPAL:	Comisión Económica Para América Latina y el Caribe
DANE:	Departamento Administrativo Nacional de Estadística
DIAN:	Dirección de Impuestos y Aduanas Nacionales
GLP:	Gas Licuado de Petróleo
IDW:	Interpolación de Distancia Inversa Ponderada
IMG:	Ingreso Mínimo Garantizado
IPM:	Índice de pobreza multidimensional
IRUI:	Índice del Registro Universal de Ingresos
EMB:	Encuesta Multipropósito de Bogotá
OCDE:	Organización para la Cooperación y el Desarrollo Económicos
PAE:	Programa de alimentación Escolar
PILA:	Planilla Integrada de Liquidación de Aportes
RA:	Registro Administrativo
RSH:	Registro Social de Hogares
RUI:	Registro Universal de Ingresos
SDH:	Secretaría Distrital de Hacienda
SDP:	Secretaría Distrital de Planeación
SISBEN:	Sistema de Identificación de Potenciales Beneficiarios de Programas Sociales
UPZ:	Unidad de Planeamiento Zonal

## INTRODUCCIÓN

El presente texto constituye el documento del Producto 4 del Proyecto que tiene como objeto analizar la estratificación como instrumento de focalización junto con otros instrumentos de política para mirar su eficiencia. Según el Anexo Técnico del Proyecto este documento está dirigido a presentar los resultados del análisis descriptivo del índice propuesto en el documento anterior (Producto 3) con relación a la estratificación, de la medición de sus errores de inclusión y exclusión, y de la segmentación de las variables e indicadores que condicionan los subsidios y contribuciones. Además, comprende la programación del software estadístico utilizado en la ejecución del proyecto y a proponer recomendaciones de política pública en torno a los temas del estudio.

En el Anexo Técnico el Producto 4 tiene el siguiente alcance:

- i. Análisis descriptivo de las variables o indicadores que intervienen en el índice, con los niveles de desagregación posibles, entre ellos: rural-urbano, localidad, clasificación por entidades, por programas, por nivel de ingresos, por tipos de cobros, entre otros.
- ii. Resultado cuantitativo del índice propuesto, con el análisis de la semaforización frente a la asignación de subsidios para la prestación de servicios públicos domiciliarios.
- iii. Relación cuantitativa del índice construido con la estratificación vigente, con la incidencia del pago de servicios públicos domiciliarios y la capacidad de pago de los hogares.
- iv. Resultado de la medición de los errores de inclusión y exclusión, comparados con la estratificación actual, y su evidencia de minimización de errores
- v. Resultado de la segmentación de las variables e indicadores que condicionan los subsidios y contribuciones en servicios públicos domiciliarios.
- vi. Programación (código) para obtener los resultados de análisis y descriptivos en software de uso libre, preferiblemente en R, con las especificaciones técnicas que se acuerden entre ambas partes, previo a la entrega de este documento.
- vii. Recomendaciones de política pública: potencial aplicación de este instrumento de declaración universal (el proceso, entidades a cargo, costos aproximados de la recolección de la información) y su eficiencia, además de los riesgos y posibles amenazas en la implementación.
- viii. Conclusiones y recomendaciones.

En las presentaciones que se realizaron los días 22 de diciembre de 2023, 14 de febrero y 8 de marzo de 2024 ante la Supervisión del contrato sobre los contenidos del Producto 4, hubo observaciones que Analytica valoró y tuvo en cuenta para los ajustes que se incorporaron en esta versión 7 del documento.

Es de advertir, como se hizo en el documento del Producto 3, que el informe se basa principalmente en los resultados de un procesamiento especial de la información de la Encuesta Multipropósito de Bogotá, 2021 (EMB 2021), a lo cual se añade una evaluación de los registros administrativos (RA) con que se contó, para contribuir a la construcción del

RUI. En primer lugar, del PILA, continuando los análisis contenidos en el documento del Producto 3. Y, en segunda instancia, de la encuesta del SISBÉN IV.

El alcance de lo logrado en relación con la utilización de los RA se vio limitado por la imposibilidad de disponer de RA que contienen información que es fundamental para el RUI, como la correspondiente a la declaración de renta de la DIAN, la de subsidios y pago de impuestos y el impedimento, basado en normas legales, sobre el manejo de información, para hacer una integración de los RA disponibles, al carecer de las claves para hacerlo. Esta limitación de la información anonimizada que se ha recibido se analiza y comenta en la parte final de este documento (Recomendaciones de política).

En los Apéndices 1 y 2 se incluyen los aspectos metodológicos de la construcción del índice con rutinas y del modelo de estimación de ingresos utilizado a partir del SISBÉN IV.

Los siguientes archivos hacen parte del Producto 4:

- C707-23 P4 Diccionario de datos (27022024).xlsx
- C707-23 P4\_diccionario\_variables\_adicionales\_emb2021 (27022024).xlsx
- C707-23 P4\_diccionario\_variables\_encuesta\_emb2021 (27022024).xlsx
- C707-23 P4 Sintaxis modelos (27022024).sps
- C707-23 P4 Programa utilizado-ubicación sintaxis-justificacion pruebas (27022024).docx
- C707-23 P4 Insumos\_Mapas.gdb (27022024).rar
- C707-23 P4 Prueba (27022024).rar
- C707-23 ANALYTICA-SDP-Propuesta diseño del RUI para Bogotá.pptx
- C707-23 P4 Anexo computacional SDP.zip

## 1. DESCRIPCIÓN DE LAS VARIABLES INVOLUCRADAS EN EL ÍNDICE

### 1.1. LOS NIVELES DE INGRESO EN EL DESARROLLO PROGRESIVO DEL RUI

El ingreso disponible ajustado es la medida de ingreso óptima, a la cual hay que dirigirse en un proceso progresivo de mediano plazo, seguida en orden secuencial por el ingreso disponible, el ingreso corriente secundario y el ingreso corriente primario. Desde una perspectiva de viabilidad progresiva del RUI, se privilegian, para su fase inicial, las categorías ingreso corriente primario e ingreso corriente secundario. En una etapa posterior sería deseable llegar a las estimaciones de ingreso disponible e ingreso disponible ajustado.

#### [1]. Ingreso Corriente Primario.

Al representar los ingresos laborales y de capital, el ingreso corriente primario sirve como indicador inicial de la capacidad de generación de ingresos del individuo. La optimización de esta variable radica en reconocer que un ingreso primario sólido proporciona una base esencial para el bienestar económico, sirviendo como punto de partida para considerar las intervenciones redistributivas. De forma óptima se tendrá información de las siguientes variables para la construcción del ingreso corriente primario:

- Ingreso por el trabajo asalariado
- Ingresos netos de personas trabajadoras por cuenta propia<sup>1</sup> y patronos<sup>2</sup>
- Ingresos de propiedad<sup>3</sup>
- Dividendos<sup>4</sup>
- Intereses<sup>5</sup>
- Alquileres<sup>6</sup>
- Ingreso total del hogar<sup>7</sup>

---

<sup>1</sup> En materia impositiva, el trabajador por cuenta propia se define como: i) la persona responsable del impuesto a las ventas del régimen común; ii) si factura ingresos brutos y sobre ellos practica retención en la fuente; iii) si los ingresos provienen en una proporción igual o superior a un ochenta por ciento (80%) de la realización de una de las actividades económicas por propia cuenta y riesgo (artículo 3 del Decreto 3032 de 2013 y Decreto 1473 de 2014).

<sup>2</sup> Los ingresos obtenidos en el desarrollo de la actividad económica del cuenta propia o patrón corresponde por lo menos al 80% del total de sus ingresos.

<sup>3</sup> Los ingresos de la propiedad son los que provienen de la posesión de bienes raíces y propiedades financieras.

<sup>4</sup> Ganancias que se obtienen de la rentabilidad de las operaciones comerciales de una compañía, durante un período determinado.

<sup>5</sup> Ingresos por el precio del dinero en un momento determinado.

<sup>6</sup> Ingresos por arrendamiento de bienes muebles o inmuebles.

<sup>7</sup> Ingresos por trabajo, más los ingresos de capital más las prestaciones y transferencias recibidas por cada uno de sus miembros, más las entradas ocasionales utilizadas en gastos del hogar.

- Ingresos por actividades profesionales independientes<sup>8</sup>
- Ingresos por regalías y derechos de Autor<sup>9</sup>

## [2]. Ingreso Corriente Secundario.

La inclusión de ingresos primarios y transferencias en efectivo y en especie en el ingreso corriente secundario permite capturar tanto los logros laborales como las ayudas sociales. Esta medida integral refleja la diversidad de fuentes de ingresos y su contribución al bienestar, considerando tanto el esfuerzo laboral como las intervenciones redistributivas. A continuación, se listan las variables que se adicionan al ingreso primario para llegar al secundario:

- Pensiones y jubilaciones
- Transferencias monetarias y en especie (subsidios) de organizaciones gubernamentales y no gubernamentales
- Otras transferencias en dinero o especie (transferencias para el sostenimiento de hijos menores, alimentos, bienes, servicios, etc.).

## [3]. Ingreso Disponible.

Al restar impuestos y contribuciones obligatorias del ingreso corriente secundario, esta variable proporciona una medida más precisa del ingreso real disponible para el individuo. La optimización de esta variable se basa en reconocer que un mayor ingreso disponible se correlaciona positivamente con un mayor bienestar social, permitiendo un mayor acceso a bienes y servicios. A continuación, se listan las variables óptimas a ser incluidas:

- Impuesto sobre la renta personal
- Impuesto sobre Sociedades
- Impuesto al Valor Agregado (IVA)
- Impuesto sobre bienes personales
- Impuesto a las ganancias de capital y rendimientos financieros.
- Impuesto a las transacciones financieras
- Impuesto sobre sucesiones y donaciones
- Contribuciones a la seguridad social
- Impuesto predial
- Impuesto a la gasolina
- Impuesto de industria y comercio (ICA)
- Otros Impuestos locales y regionales

---

<sup>8</sup> Ingresos por prestación de servicios profesionales de manera independiente.

<sup>9</sup> Ingresos que perciben escritores, fotógrafos, traductores, pintores, ilustradores, etc.

[4]. Ingreso disponible ajustado.

La inclusión de transferencias sociales en el ingreso disponible ampliado refleja directamente el impacto de las políticas de redistribución de ingresos en el bienestar social. Las transferencias sociales, al proveer recursos adicionales a aquellos con ingresos más bajos, contribuyen significativamente a la mejora del bienestar. A continuación, se listan las variables óptimas a ser incluidas:

- Transferencias monetarias gubernamentales.
- Costos de los programas sociales recibidos.
- Subsidios sociales (por la prestación de servicios)
- Becas y ayudas educativas no monetarias
- Programas de asistencia Social
- Subvenciones no monetarias por desempleo
- Ayudas por discapacidad
- Subsidios para la vivienda
- Créditos fiscales
- Deducciones fiscales
- Programas de redistribución de renta
- Subsidios a la salud
- Beneficios tributarios por ingresos bajos
- Incentivos para la educación
- Programas de alivio de deuda
- Subsidios energéticos
- Otros programas de transferencias y redistribución

La optimalidad de estas variables se fundamenta en su capacidad para abarcar tanto las dimensiones laborales como las redistributivas del ingreso y las transferencias sociales. Al considerar los subcomponentes de cada variable, se logra una evaluación exhaustiva de la situación financiera de un individuo, brindando así una perspectiva más completa del bienestar social y permitiendo una medición más precisa de los efectos de las políticas económicas y sociales en la sociedad.

Se propone que el ingreso corriente secundario, y no el primario, sea el referente central para la construcción inicial del RUI. El RUI es un instrumento de focalización, pero parte de un proceso donde existen subsidios y transferencias públicas de distinto tipo. Para orientar el gasto público asignado a los programas sociales desde el RUI, de una manera articulada y dinámica es indispensable contar desde su inicio y en sus etapas futuras con la información de las transferencias sociales<sup>10</sup>. Tomarlas en cuenta no constituye, como podría

---

<sup>10</sup> Congreso de la República (2001). Ley 715 de 2001, Por la cual se dictan normas orgánicas en materia de recursos y competencias de conformidad con los artículos 151, 288, 356 y 357 (Acto Legislativo 01 de 2001) de la Constitución Política y se dictan otras disposiciones para organizar la prestación de los servicios de educación y salud, entre otros. Consultado el 12 de enero de 2024 de: [http://www.secretariasenado.gov.co/senado/basedoc/ley\\_0715\\_2001.html](http://www.secretariasenado.gov.co/senado/basedoc/ley_0715_2001.html). Según el artículo 94 de esta ley,

pensarse desde una visión estática, una “endogeneidad”, en el sentido de que la focalización se hace siguiendo un orden que va del ingreso primario, para después hacer la asignación de los subsidios y transferencias sobre esta base, sin tomar en cuenta los subsidios actuales. Para evitar esta “endogeneidad” se requeriría, bajo tal visión, asignar los subsidios únicamente con base en el ingreso primario. En un enfoque dinámico, que corresponda a los requerimientos prácticos de la orientación de las políticas públicas, la focalización se lleva a cabo en un proceso donde ya hay una asignación dada de subsidios que busca ser optimizada. Por eso se requiere conocer esa asignación para mejorarla. El proceso que vivió Bogotá con la pandemia y la puesta en marcha de la estrategia del Ingreso Mínimo Garantizado (IMG), tomando en cuenta la distribución existente de subsidios de los distintos programas, para mejorarla y completarla, es una expresión práctica de cómo el ingreso primario no es suficiente para orientar la puesta en marcha del RUI .

## 1.2. LOS COMPONENTES DEL ÍNDICE

Se describen en esta sección las variables principales que intervienen en el índice de ingresos propuesto. De esta forma, se busca dar una visión sobre sus componentes, y aportar elementos que ayuden a definir la vía para su construcción visualizando prioridades desde un punto de vista práctico.

En primer lugar, se hace referencia al peso que tienen esas variables por tipo de aportante de ingreso. A continuación, se analiza en qué magnitud intervienen para dar cuenta de los ingresos personales y de los hogares para el conjunto de la población.

El análisis se hace con base en la información de la EMB 2021, tomando en cuenta una variable que expresa, en forma aproximada, el ingreso corriente secundario a nivel de persona, y seis variables que intervienen significativamente en el índice. Se lleva a cabo para la población de 12 y más años (en edad de trabajar) ya que la encuesta no registra ingresos para los menores de esa edad.

Las variables consideradas son<sup>11</sup>:

- Ingreso total a nivel personas (N\_INGTOT\_PER\_): puede considerarse un “proxy” del Ingreso personal corriente secundario, incorporando lo investigado en la

---

se entiende por focalización “...el proceso mediante el cual se garantiza que el gasto social se asigne a los grupos de población más pobre y vulnerable.” El artículo 1º del decreto 4816 señala que “Los instrumentos de focalización del gasto social son herramientas estadísticas y técnicas que permiten identificar y clasificar los potenciales beneficiarios de los programas de gasto social.”

<sup>11</sup> Las variables se encuentran en la base de datos de la EMB 2021. DANE (2022). EMB 2021 <https://www.dane.gov.co/index.php/estadisticas-por-tema/pobreza-y-condiciones-de-vida/encuesta-multiproposito>. Información actualizada el 01 de julio de 2022. Los nombres de las variables en la base de datos se colocan entre paréntesis.

encuesta sobre ingresos laborales, transferencias por pensiones o jubilaciones y por ayudas<sup>12</sup>. e ingresos por arriendos<sup>13</sup> (capital).

- Ingreso laboral actividad principal<sup>14</sup> (N\_IMPAES\_): Ingreso monetario de primera actividad imputado (sólo para faltantes, extremos o ceros inconsistentes).
- Ingreso laboral actividad secundaria N\_ISAES\_: Ingreso de segunda actividad imputado (sólo para faltantes, extremos o ceros inconsistentes).
- Ingreso laboral monetario de desocupados e inactivos N\_IMDIES\_: Ingreso monetario de desocupados e inactivos imputado (sólo para faltantes, extremos o ceros inconsistentes).
- Ingreso por Jubilaciones o pensiones N\_IOF2ES\_: Ingreso por jubilaciones o pensiones imputado (sólo para faltantes, extremos o ceros inconsistentes).
- Ingreso por ayudas N\_IOF3ES\_: Ingreso por ayudas imputado (sólo para faltantes, extremos o ceros inconsistentes).
- Ingreso por arriendos (N\_IOF6ES\_: Ingreso por arriendos imputado (sólo para faltantes, extremos o ceros inconsistentes).

### 1.2.1. Los perceptores por tipo de ingreso

En los gráficos 1 y 2 se encuentran representados el número de perceptores<sup>15</sup> de cada variable y el valor promedio que les corresponde. Se observa que el número de perceptores se concentra en los ingresos laborales, encontrándose que en los otros tipos de ingresos participa una porción minoritaria. Es así, por ejemplo, que los aportantes a ingresos laborales en la principal actividad superan los 4 millones 300 mil trabajadores, mientras que los individuos que reciben arriendos son menos de 200.000.

En contraste, los ingresos promedio de los jubilados o pensionados son los más elevados, a pesar de representar una categoría con poca participación poblacional. También los arriendos medios son altos en contraposición con su peso poblacional.

---

<sup>12</sup> Entre otros, los apoyos a la educación, la alimentación, la salud, provenientes de programas sociales gubernamentales.

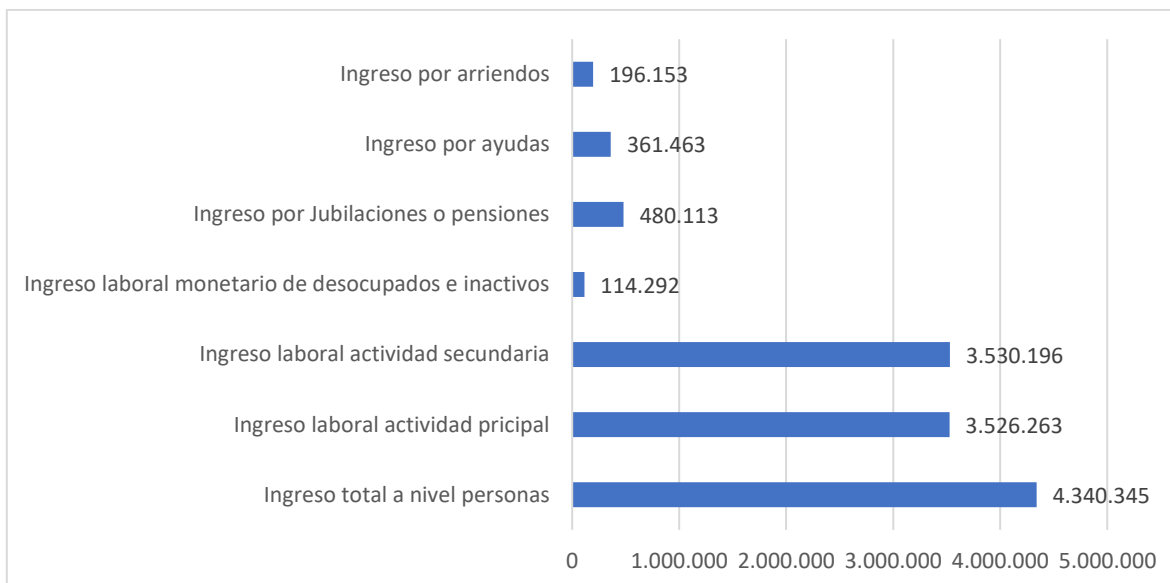
<sup>13</sup> Valor recibido por concepto de arrendamiento de bienes muebles e inmuebles. Dirección de Impuestos y Aduanas Nacionales -DIAN- (2024). Formulario 210 2024. Consultado el 17 de enero de 2024 de: [https://www.dian.gov.co/atencionciudadano/formulariosinstructivos/Formularios/2024/Formulario\\_210\\_2\\_024.pdf](https://www.dian.gov.co/atencionciudadano/formulariosinstructivos/Formularios/2024/Formulario_210_2_024.pdf)

<sup>14</sup> Actividad económica que genera el mayor valor de ingresos en el periodo gravable. Dirección de Impuestos y Aduanas Nacionales -DIAN- (2024). Formulario 210 2024. Consultado el 17 de enero de 2024 de: [https://www.dian.gov.co/atencionciudadano/formulariosinstructivos/Formularios/2024/Formulario\\_210\\_2\\_024.pdf](https://www.dian.gov.co/atencionciudadano/formulariosinstructivos/Formularios/2024/Formulario_210_2_024.pdf)

<sup>15</sup> Quienes registran respuesta en la encuesta.

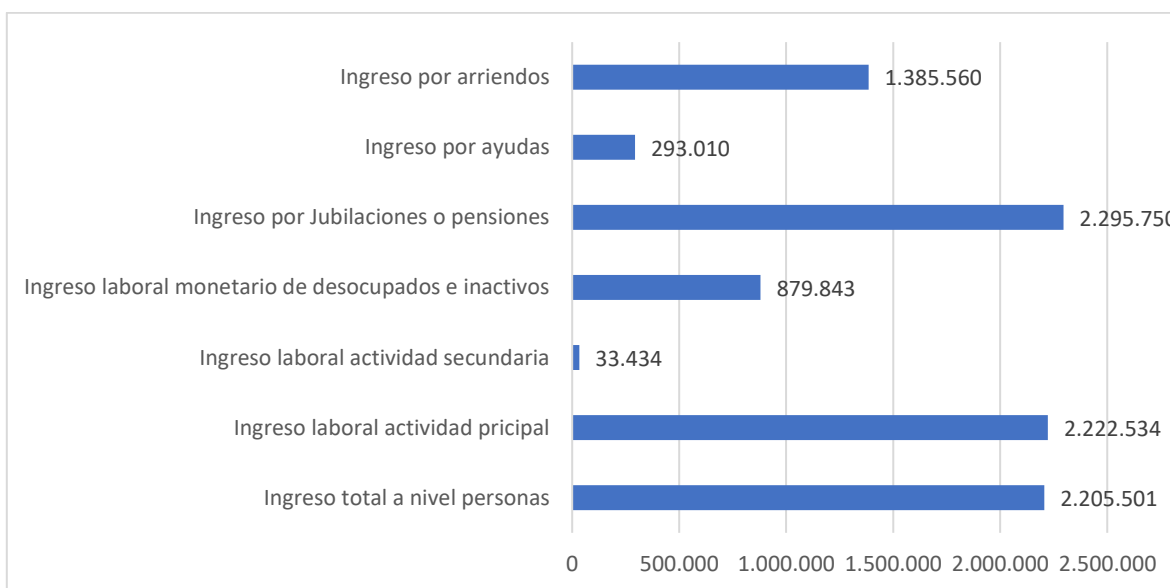


*Gráfico 1. Perceptores de ingresos de 12 y más años por tipo de ingreso- Bogotá-2021*



*Fuente: Procesamiento propio de la EMB 2021.*

*Gráfico 2. Promedio mensual de los tipos de ingresos de personas de 12 y más años por tipo de ingreso- Bogotá-2021 (en pesos corrientes)*

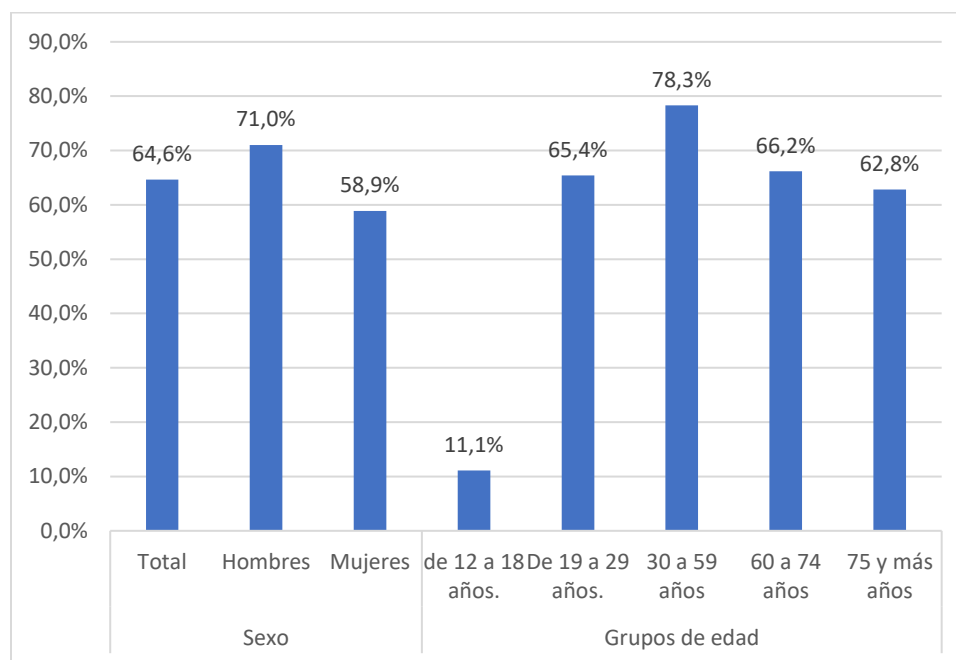


*Fuente: Procesamiento propio de la EMB 2021.*

El porcentaje de perceptores cambia, además, con las características de los grupos poblacionales. Es apreciablemente mayor en los hombres (71%), que en las mujeres (59%),

y aumenta con la edad hasta los 60 años, para disminuir entre los adultos mayores (Gráfico 3). Se destaca, además, que casi una tercera parte de la llamada población en edad de trabajar no tenga ingresos.

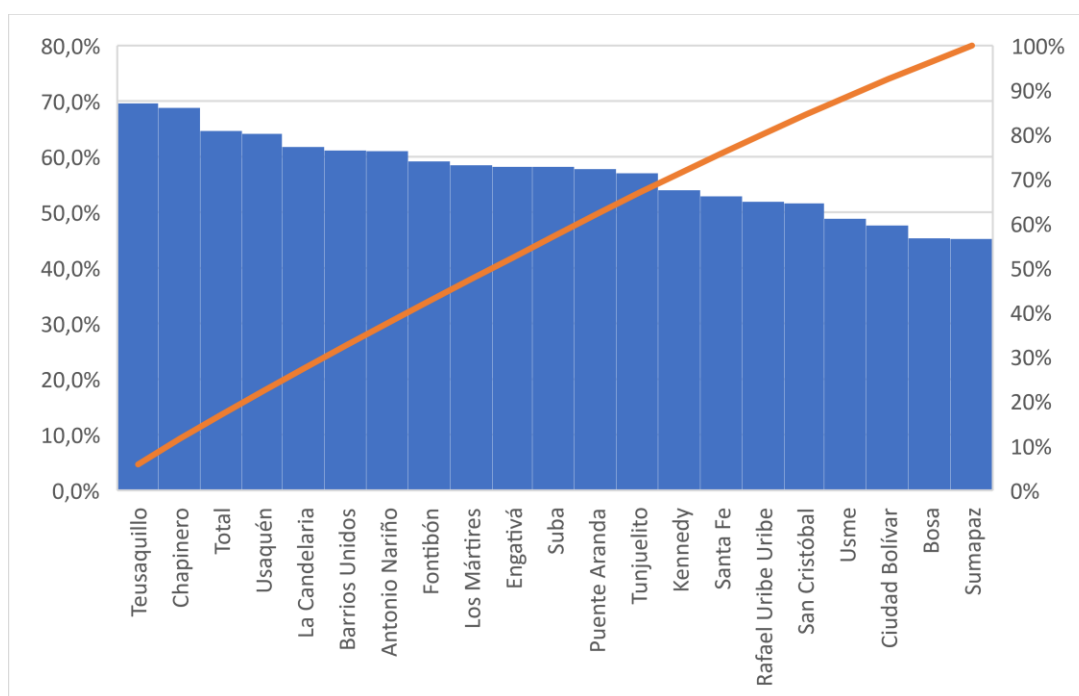
*Gráfico 3. Porcentaje de perceptores de ingresos, por sexo y grupos de edad- Bogotá- 2021*



*Fuente: Procesamiento propio de la EMB 2021.*

Se presentan igualmente diferencias en la participación de los perceptores entre las localidades de Bogotá. Tiende a aumentar en concordancia con las condiciones más favorables de vida. Como en otras problemáticas, las Localidades del Centro-Norte y Norte Teusaquillo, Usaquén y Chapinero, a las que se une La Candelaria, tienen las mayores tasas de participación, frente a las localidades del Sur que, junto con más altos niveles de pobreza, exhiben las menores. En este resultado incide la composición etaria más vieja de las localidades con más favorables condiciones de vida (Gráfico 4).

Gráfico 4. Porcentaje de perceptores de ingresos de 12 y más años por Localidad-Bogotá-2021

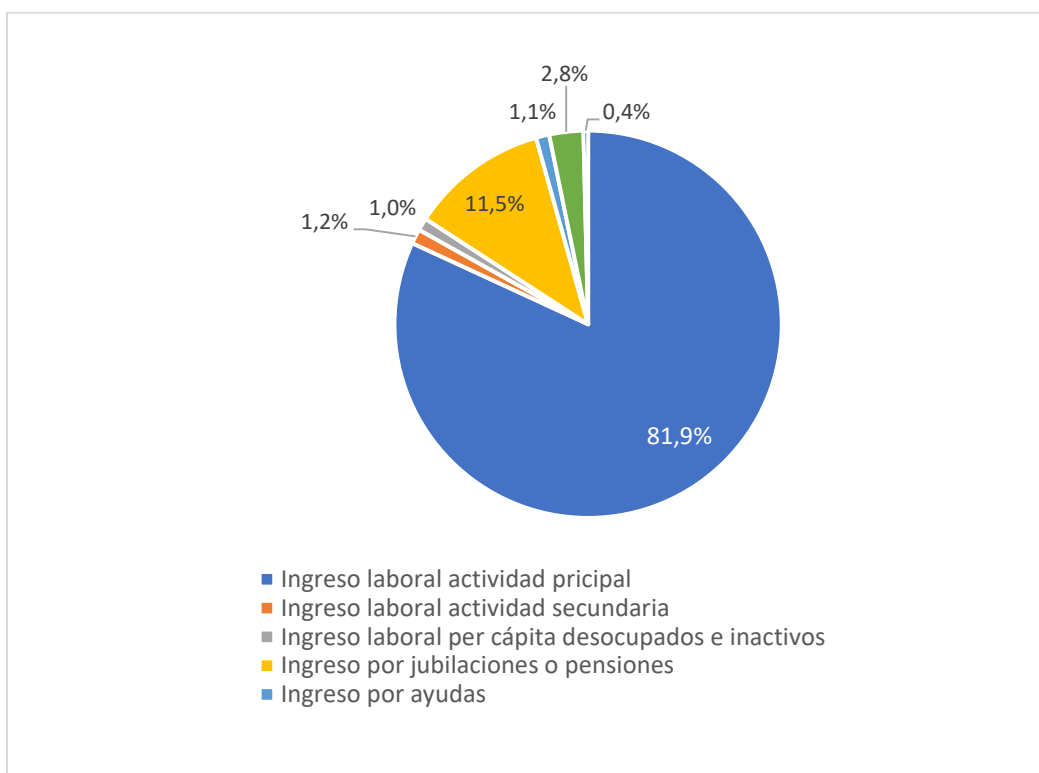


Fuente: Procesamiento propio de la EMB 2021.

#### 1.2.2. Diferencias en los ingresos por grupos poblacionales y en el aporte de los distintos tipos de ingresos

Los seis tipos de ingresos parciales considerados representan, en su conjunto, una proporción superior al 99% del total de los ingresos secundarios, como se ilustra en el Gráfico 5. El 84.2% de estos ingresos provienen de los ingresos laborales y el 11.5% de las transferencias de pensiones o jubilaciones. Los otros rubros aportan el 6.5% restante.

Gráfico 5. Participación de los distintos tipos de ingreso en el ingreso secundario total personal- Bogotá-2021



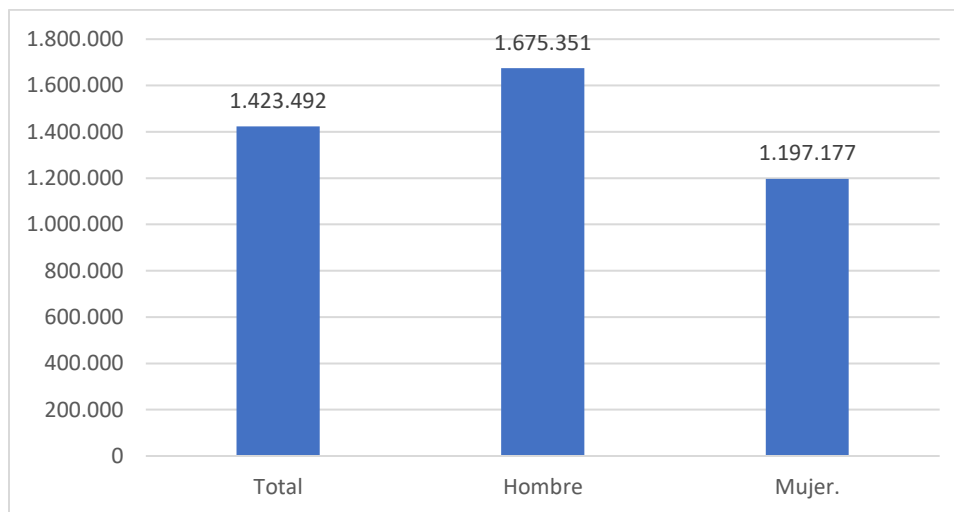
Fuente: Procesamiento propio de la EMB 2021.

El promedio por persona de este ingreso para 2021 era de \$1.423.492. Los montos y la composición de los ingresos tienen, no obstante, amplia diversidad dentro de segmentos poblacionales distinguidos por distintas características.

El ingreso secundario de las mujeres es inferior al de los hombres en un 28.5% (Gráfico 6). Es una diferencia apreciable que está afectada por la menor participación laboral femenina y por sus niveles de remuneración más bajos. Las diferencias por sexo, que marcan condiciones menos favorables para las mujeres, se expresan recurrentemente en distintos indicadores.

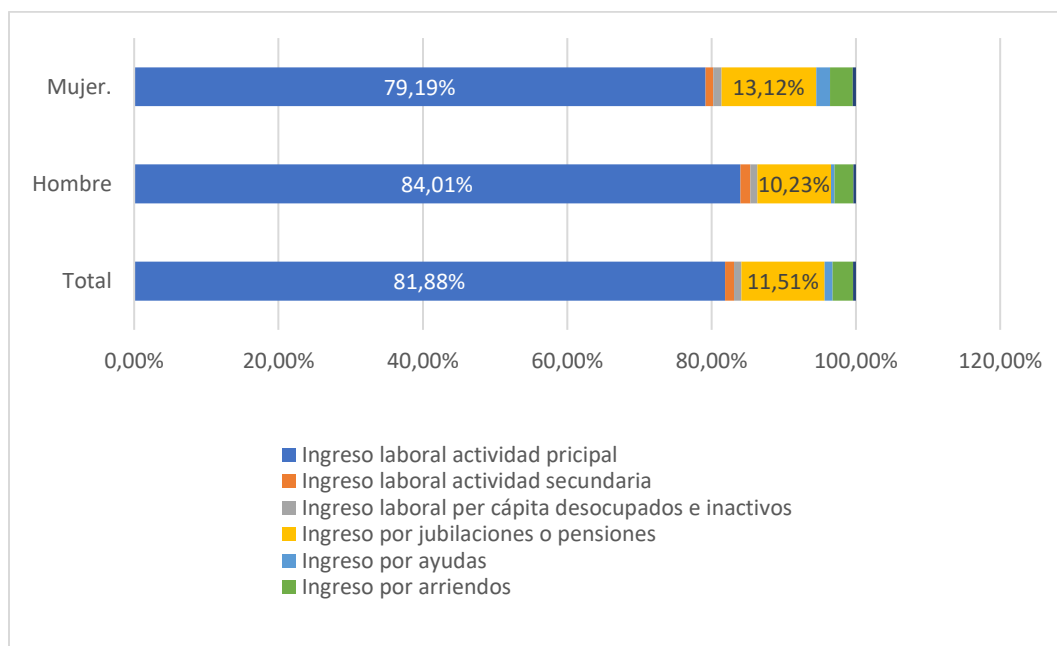
La diversidad en la composición sociolaboral por sexo se manifiesta también en diferencias en las participaciones de los distintos tipos de ingreso en los ingresos totales (Gráfico 7). Entre los ingresos de las mujeres es menor la participación de los ingresos laborales y mayor la de jubilaciones y pensiones. Igualmente es más elevada en su ingreso el aporte de los arriendos y las ayudas.

*Gráfico 6. Ingreso secundario total, personal, por sexo-Bogotá-2021*



*Fuente: Procesamiento propio de la EMB 2021.*

*Gráfico 7. Participación de los distintos tipos de ingreso en el ingreso secundario total personal, por sexo- Bogotá-2021*



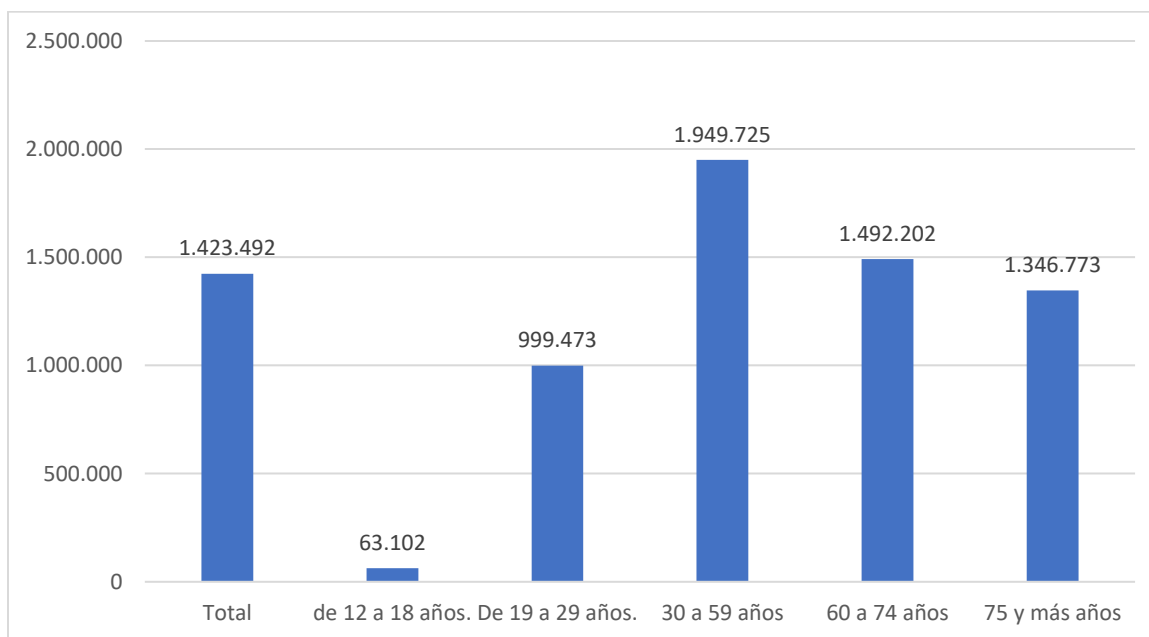
*Fuente: Procesamiento propio de la EMB 2021.*

Son más contrastadas las distancias en los ingresos por grupos de edad (Gráfico 8), que reflejan las etapas del ciclo de vida laboral y que, del mismo modo, se expresan en varios indicadores. Los menores de 19 años tienen ingresos reducidos en su fase de incorporación

al mercado laboral y de estudios secundarios y superiores. Aumentan los ingresos en las edades posteriores para decaer en las que se ubican los adultos mayores.

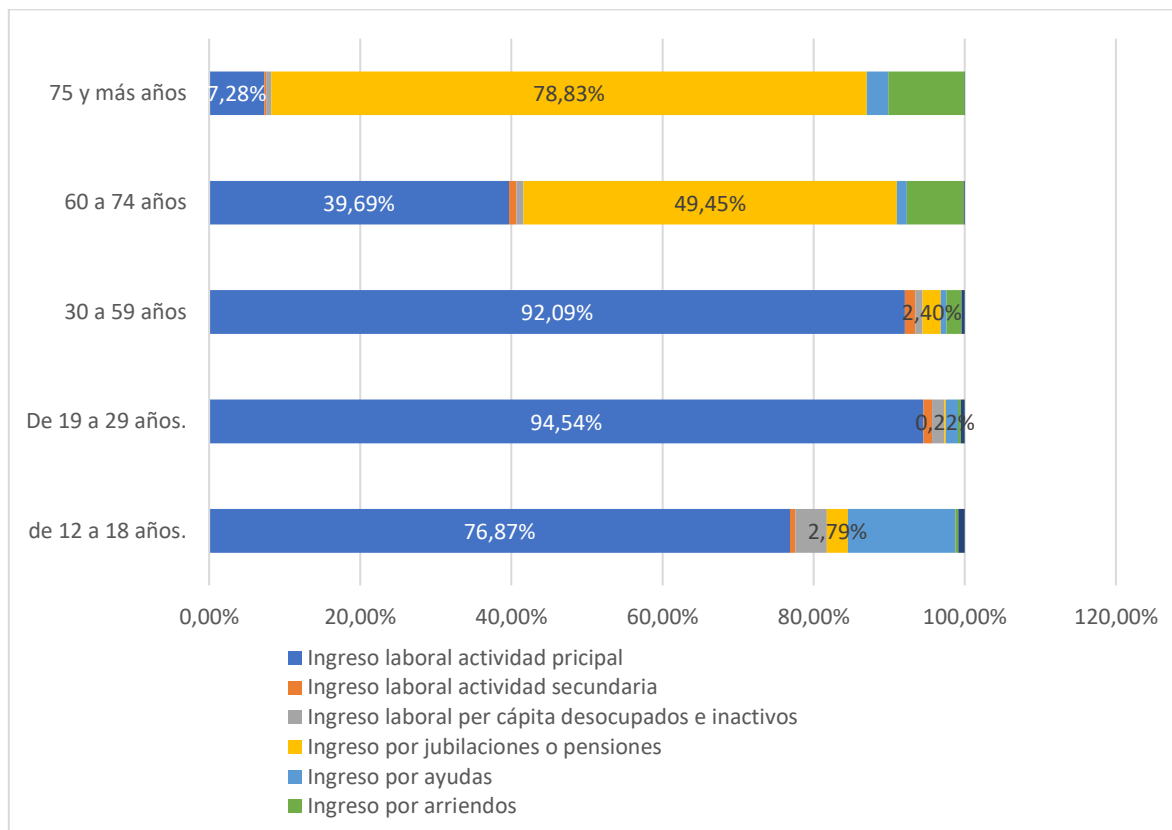
Los cambios en la participación de los distintos tipos de ingreso en el ingreso total son también acordes con el recorrido del ciclo vital (Gráfico 9). Entre los adolescentes, sus escasos ingresos tienen origen principalmente en el trabajo, pero juegan un papel destacado también las ayudas. En las etapas siguientes, hasta los 60 años, los ingresos laborales representan más del 90% de los ingresos, para entrar a jugar un papel destacado los ingresos por pensiones y jubilaciones a partir de esta edad, Para quienes se encuentran entre los 60 y 74 años las pensiones y jubilaciones contribuyen con cerca a la mitad de sus ingresos y para los que pasan este tramo, una proporción cercana al 80%.

*Gráfico 8. Ingreso secundario total, personal, por grupos de edad-Bogotá-2021 (pesos corrientes)*



*Fuente: Procesamiento propio de la EMB 2021.*

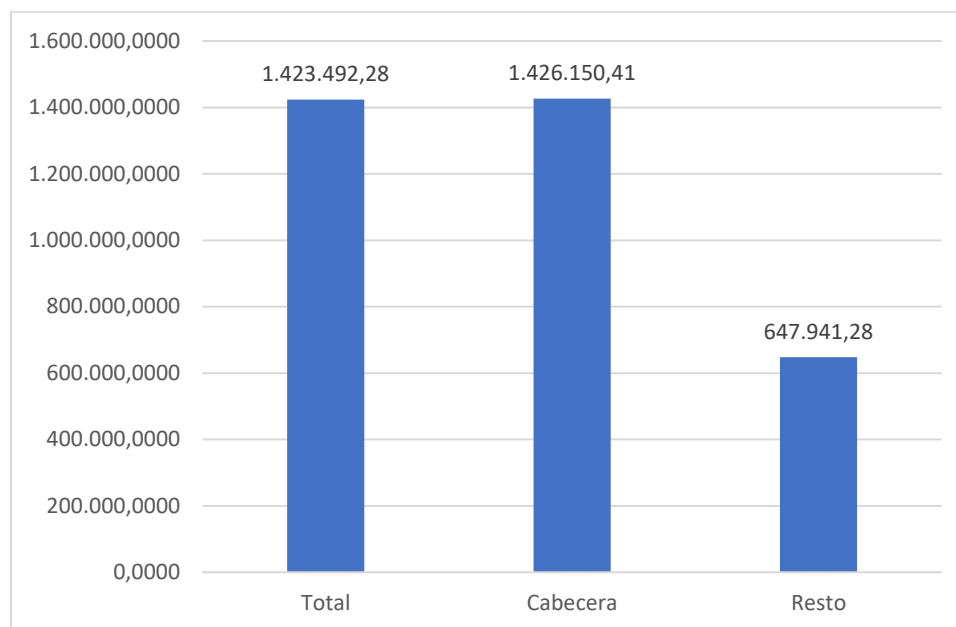
*Gráfico 9. Participación porcentual de los distintos tipos de ingreso en el ingreso secundario total personal, por grupos de edad- Bogotá-2021*



*Fuente: Procesamiento propio de la EMB 2021.*

Para el área rural el ingreso per cápita es de \$647.041, 2.2 veces menor al de la cabecera municipal. Así se manifiesta la crónica desventaja de las poblaciones campesinas, incluso en las proximidades de una gran ciudad (Gráfico 10).

*Gráfico 10. Participación de los distintos tipos de ingreso en el ingreso secundario total personal, por zonas urbana (cabecera) y rural (resto)- Bogotá-2021*



*Fuente: Procesamiento propio de la EMB 2021.*

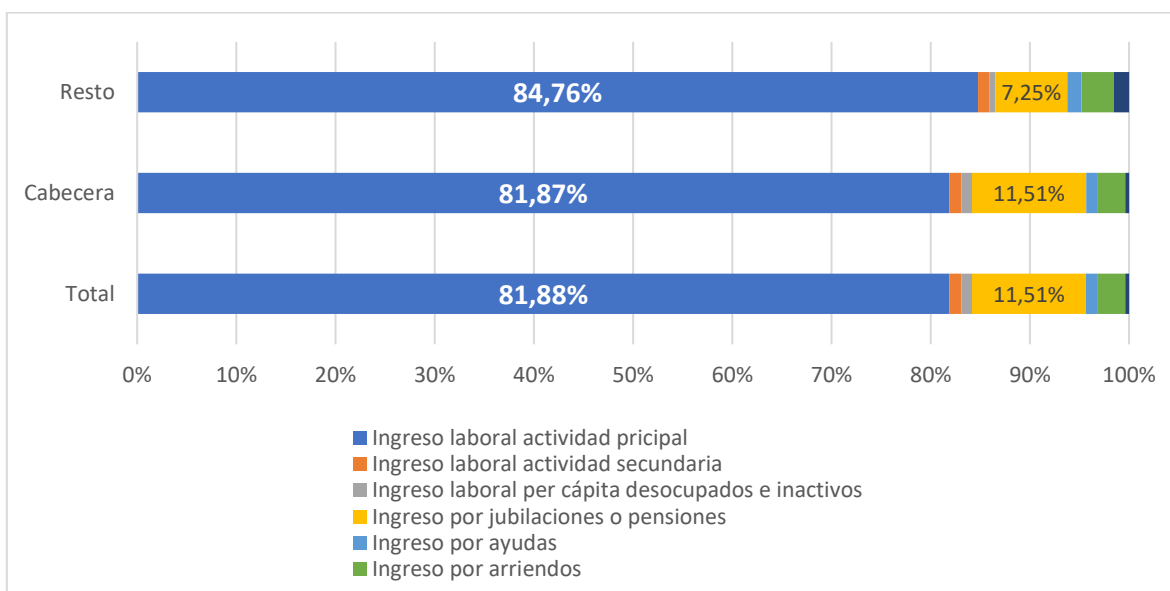
Esta situación se capta asimismo en las considerables diferencias en los promedios per cápita de ingresos entre las Localidades (Gráfico 12). En la rural de Sumapaz se perciben promedios de ingresos 10.6 veces inferiores a los de la Localidad de Chapinero, que exhibe los niveles más altos al lado de las de Teusaquillo y Usaquén. Se pone de manifiesto así la alta desigualdad en la distribución del ingreso a nivel territorial, que trasciende la contraposición campo-ciudad.

Esta situación guarda correspondencia con el peso que tiene cada tipo de ingreso en los ingresos totales. Se constata, como tendencia, que en las zonas rurales (resto) como en Sumapaz, es superior la participación de los ingresos laborales, y a medida que tienen una mayor participación en el total, los montos promedio de los ingresos son más bajos (Gráficos 11 y 13).

Los ingresos por pensiones y jubilaciones, que captan el 11.5% del ingreso total de Bogotá, y los ingresos por arriendos, con 2.8%, tienen participación mayor en las Localidades con más elevados ingresos. Las ayudas, que incluyen los subsidios de programas estatales, tienen en la encuesta un aporte relativamente pequeño: 1.1%, llegando a 1.4% en la zona rural y a un máximo de 2.2% en la Localidad de La Candelaria.

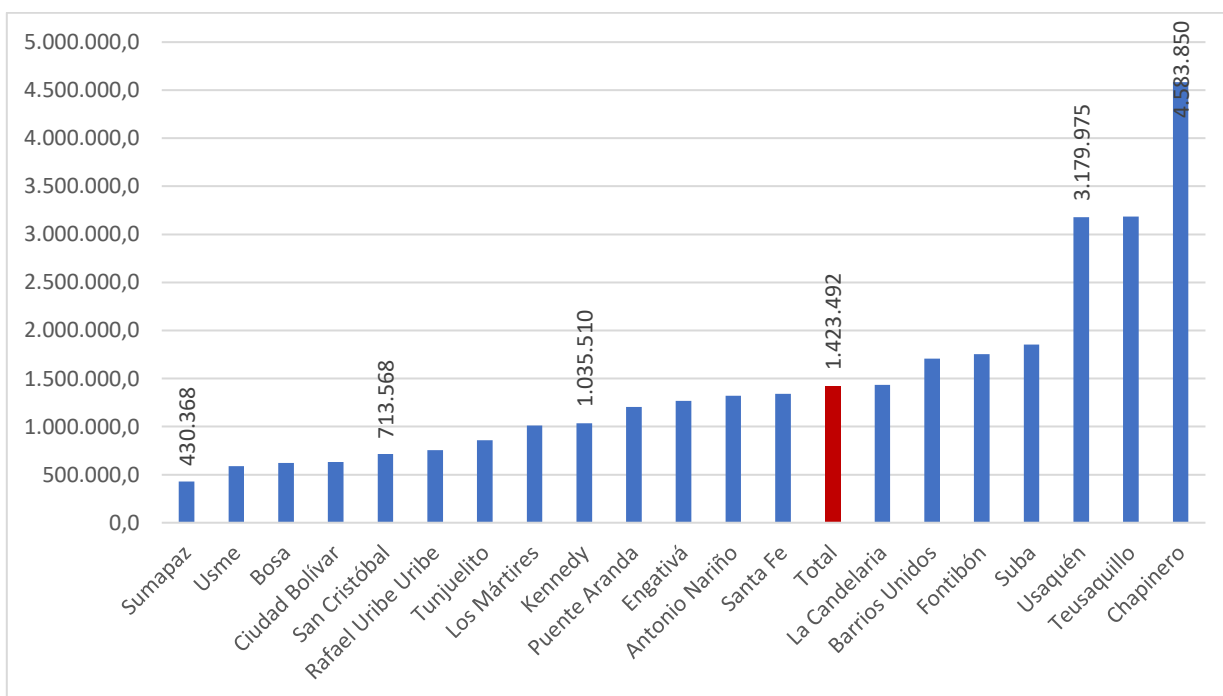


Gráfico 11. Participación de los distintos tipos de ingreso personal en el ingreso total- Bogotá-2021



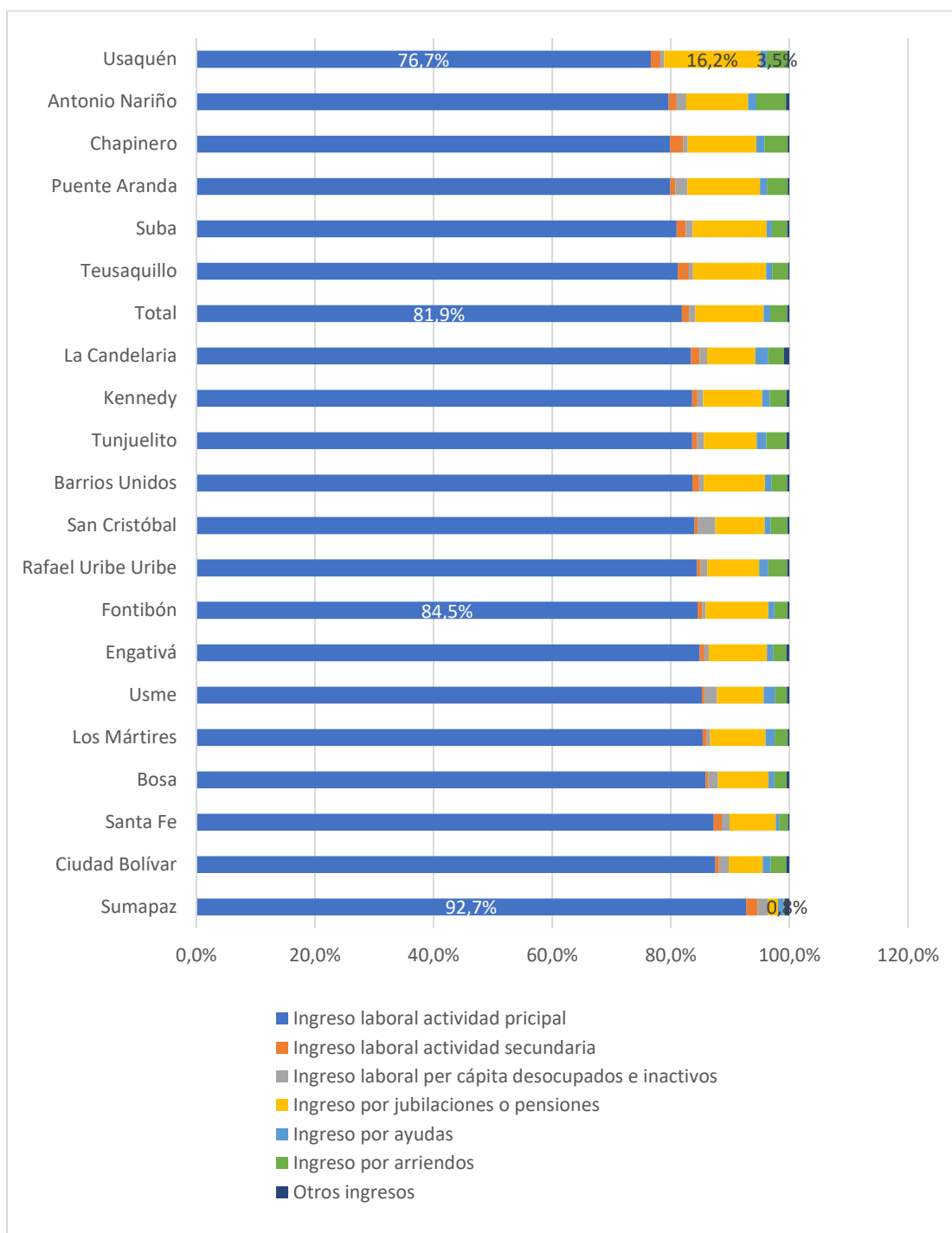
Fuente: Procesamiento propio de la EMB 2021.

Gráfico 12. Ingreso secundario total, personal, por Localidad-Bogotá-2021



Fuente: Procesamiento propio de la EMB 2021.

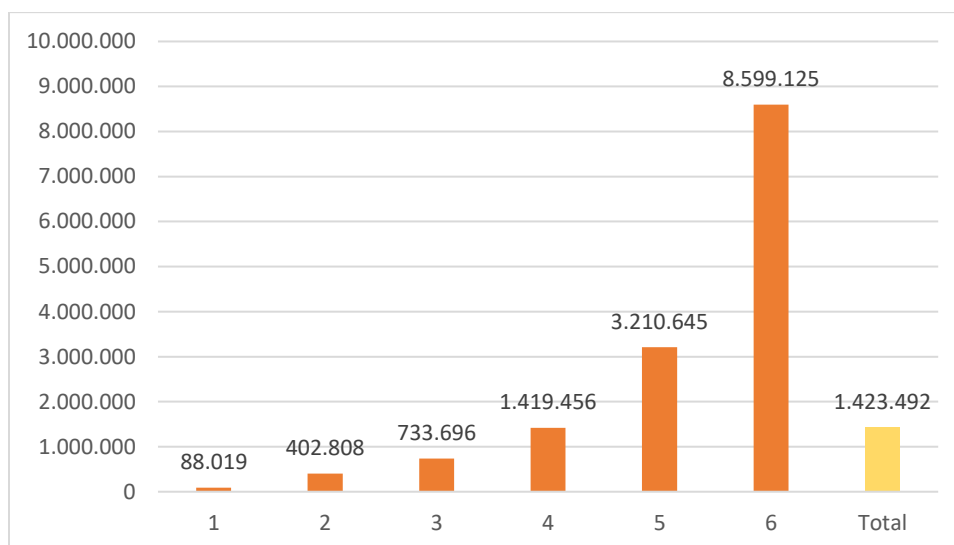
Gráfico 13. Participación de los distintos tipos de ingreso personal en el ingreso total, por Localidad- Bogotá-2021



Fuente: Procesamiento propio de la EMB 2021.

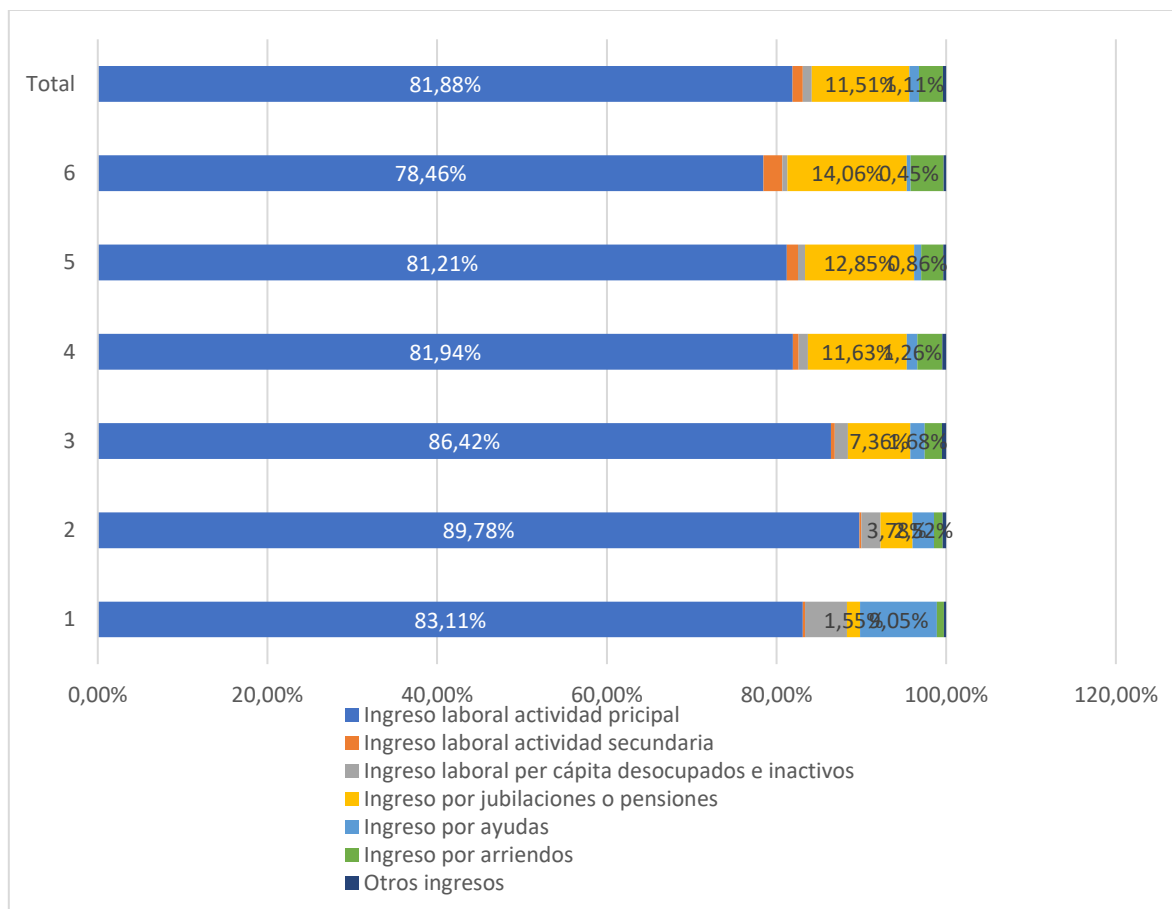
Entre los grupos poblacionales definidos en el modelo 2 del índice de ingresos propuesto se encuentran grandes diferencias en sus ingresos per cápita. El ingreso del primer grupo representa apenas el 1% del ingreso del grupo 6, y el del grupo 2 el 4.7% del mismo. Y el ingreso del grupo 5 no llega a ser la mitad del correspondiente al grupo superior (Gráfico 14). Estas diferencias se encuentran relacionadas con las de las participaciones de los tipos de ingresos en el total. A medida que el grupo es más alto, baja la participación de los ingresos laborales y asciende la de los arriendos y las pensiones (Gráfico 15).

*Gráfico 14. Ingreso per cápita secundario , por grupos del modelo 2 del índice de ingresos-Bogotá-2021 (pesos corrientes)*



*Fuente: Procesamiento propio de la EMB 2021.*

*Gráfico 15. Participación porcentual de los distintos tipos de ingreso en el ingreso secundario total personal, por grupos del modelo 2 del índice de ingresos- Bogotá-2021*



*Fuente: Procesamiento propio de la EMB 2021.*

Es de interés comparar los niveles de ingreso y su composición, entre los grupos de población que reciben subsidios a través de los principales programas públicos, y los que no participan en ellos. Se obtiene de esta forma un acercamiento a la forma como opera la focalización en estos programas y como se ve afectada la distribución de los ingresos al participar en ellos, poniendo a prueba por adelantado, el tipo de análisis que permitiría el RUI en relación con la evaluación del impacto de las intervenciones públicas.

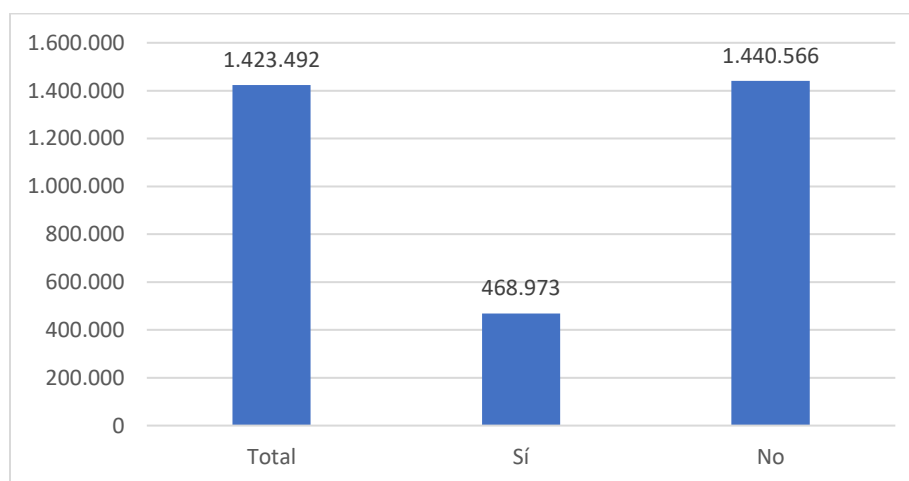
En este ejercicio se analizan los programas de Familias en Acción y de Adultos mayores, que son investigados en la encuesta. La población participante que estima la encuesta en estos programas es de 118.359 para el primero, y de 183.791 para el segundo, recibiendo subsidios promedios mensuales por persona de \$12.185 y de \$15.255 respectivamente.

Según los gráficos 16 y 18, los ingresos promedio por persona de quienes reciben los subsidios de los dos programas son notablemente inferiores a la de quienes no reciben

asistencia de éste. Los beneficiarios de Familias en Acción tienen, en promedio, ingresos 3 veces más bajos que los no beneficiarios y en el programa de Adulto Mayor es 3,8 veces inferior. Son medidas que van en la vía de sustentar *grosso modo* un proceso de focalización bien encaminado.

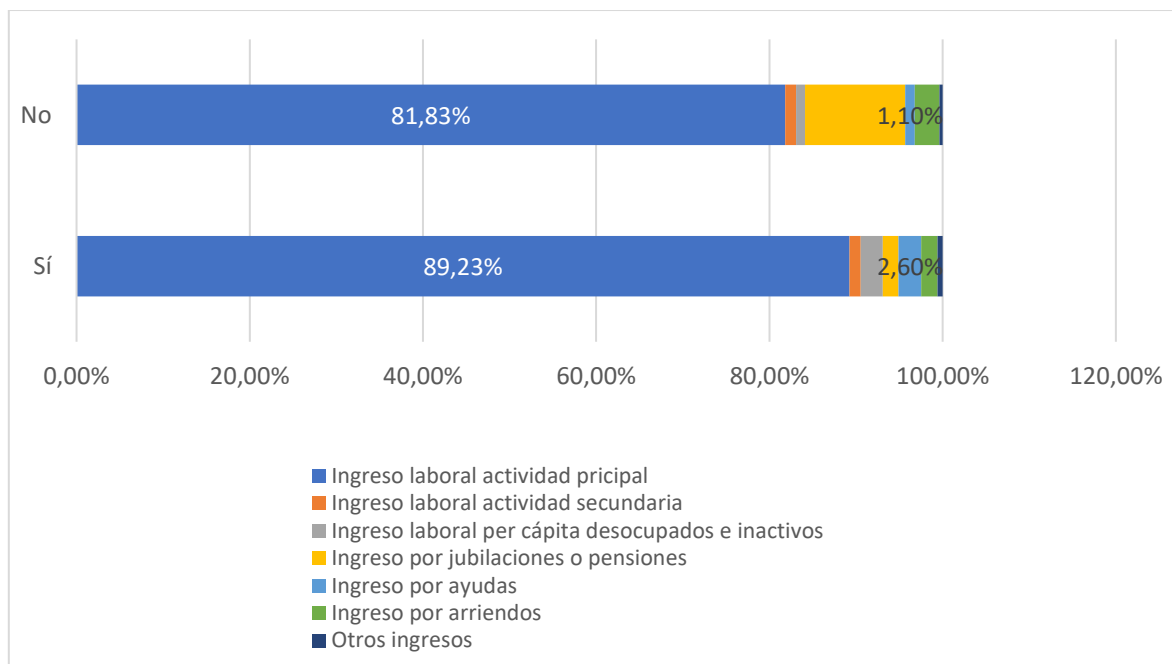
Por su parte, los gráficos 17 y 19, muestran cómo la composición de los ingresos de los dos grupos difiere, en tanto la de las poblaciones cubiertas tienen un mayor aporte de los ingresos laborales y menores de los ingresos provenientes de pensiones y ahorros. En el caso de Familias en Acción los apoyos que reciben suman 2,6% de los ingresos, mientras que los que no participan en los programas reciben apoyos que representan el 1,1% de sus ingresos. Y para el caso del programa de Adulto Mayor estos mismos porcentajes llegan a 4,1% y 1,1%. Sobre esta base se confirma que los dos programas tienen un efecto redistributivo, que disminuye la concentración de los ingresos primarios.

*Gráfico 16. Ingreso secundario total, personal, por participación en el programa de Familias en Acción- -Bogotá-2021 (pesos corrientes)*



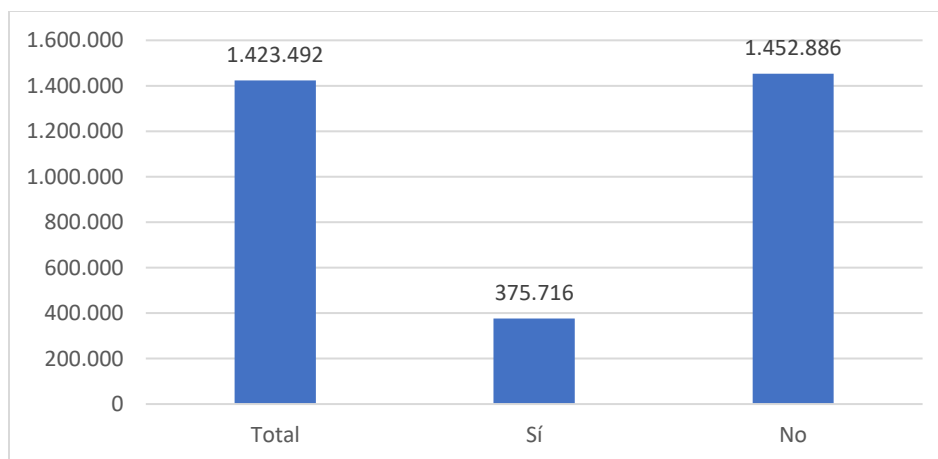
*Fuente: Procesamiento propio de la EMB 2021.*

*Gráfico 17. Participación porcentual de los distintos tipos de ingreso en el ingreso secundario total personal, por participación en el programa de Familias en Acción-Bogotá-2021*



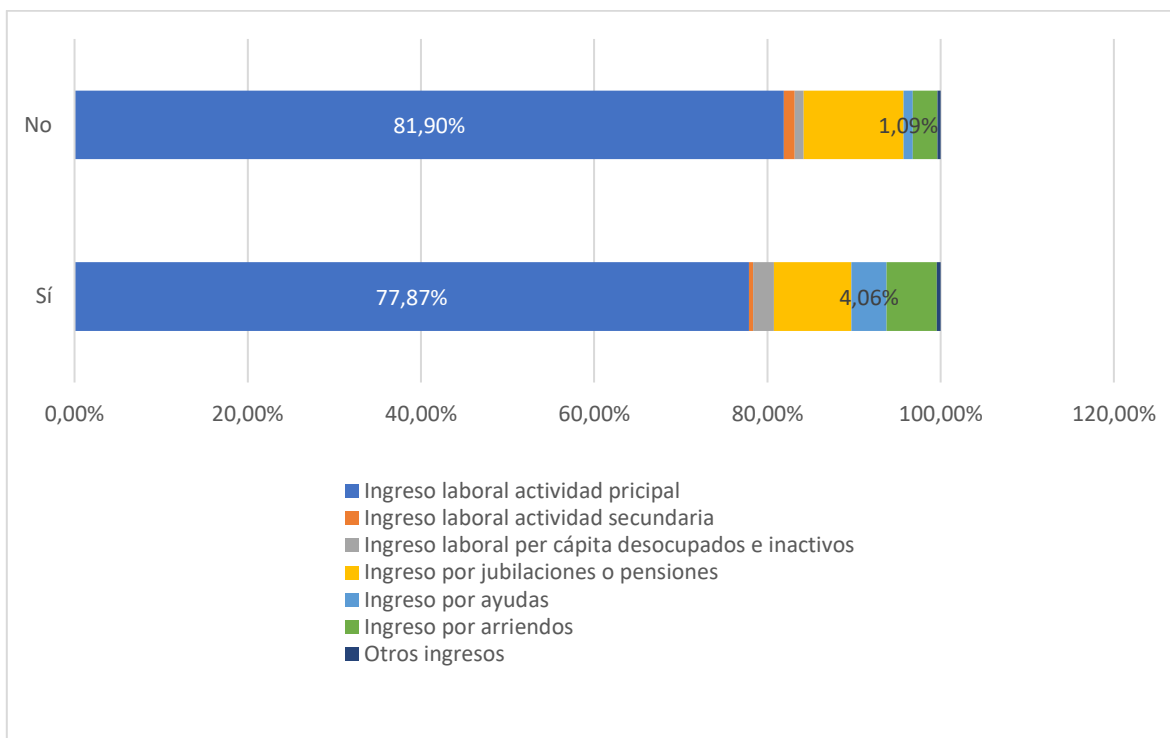
*Fuente: Procesamiento propio de la EMB 2021.*

*Gráfico 18. Ingreso secundario total, personal, por participación en el programa de Adultos Mayores-Bogotá-2021 (pesos corrientes)*



*Fuente: Procesamiento propio de la EMB 2021.*

*Gráfico 19. Participación porcentual de los distintos tipos de ingreso en el ingreso secundario total personal, por participación en el programa de Adultos Mayores- Bogotá- 2021*



*Fuente: Procesamiento propio de la EMB 2021.*

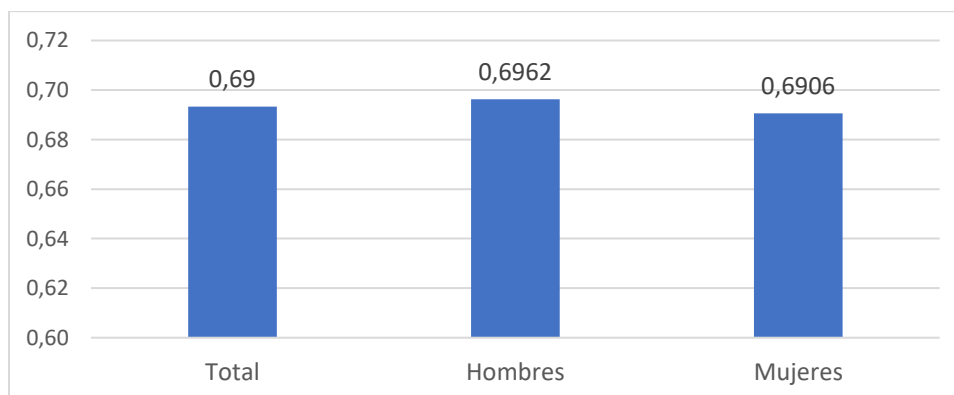
## 2. RESULTADO CUANTITATIVO DEL ÍNDICE PROPUESTO, CON EL ANÁLISIS DE LA SEMAFORIZACIÓN FRENTE A LA ASIGNACIÓN DE SUBSIDIOS PARA LA PRESTACIÓN DE SERVICIOS PÚBLICOS DOMICILIARIOS

### 2.1 Resultado cuantitativo del índice propuesto.

Se exponen algunos resultados de la estimación del índice RUI de ingresos secundarios por algunas variables socio demográficas diferenciadoras. El propósito de hacerlo es aportar elementos que apoyen el análisis crítico de la validez del índice en su aplicación práctica.

Se inicia con el examen del índice promedio por sexo (Gráfico 20), que muestra niveles muy similares para mujeres y hombres, en torno a 0,69. Este resultado evidencia un alto grado de confluencia en la ubicación de ambos sexos en los valores del índice. No expresa, sin embargo, las considerables desigualdades en los ingresos entre mujeres y hombres, que fueron tratadas anteriormente tomando en cuenta los diferenciales en los ingresos individuales, en vez de los niveles de ingreso de las unidades de gasto.

*Gráfico 20. Índice promedio RUI por sexo- Bogotá-2021*

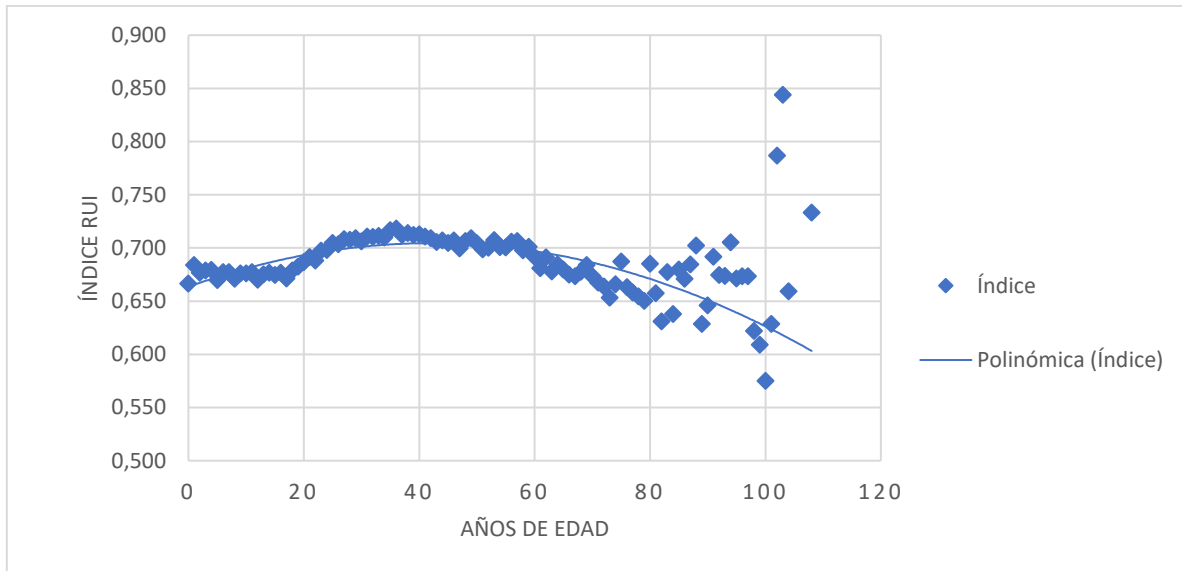


*Fuente: Procesamiento propio de la EMB 2021.*

En los índices promedios por edades simples, exhibidos en el Gráfico 21, se encuentra la tendencia a obtener índices más elevados a medida que la edad avanza, hasta cerca de los 70 años, cuando empiezan a disminuir. Es un comportamiento coherente con las etapas del ciclo de vida y especialmente con las fases de la actividad económica y la vida laboral. Además, los hogares con ingresos más bajos tienden a tener una composición más joven.



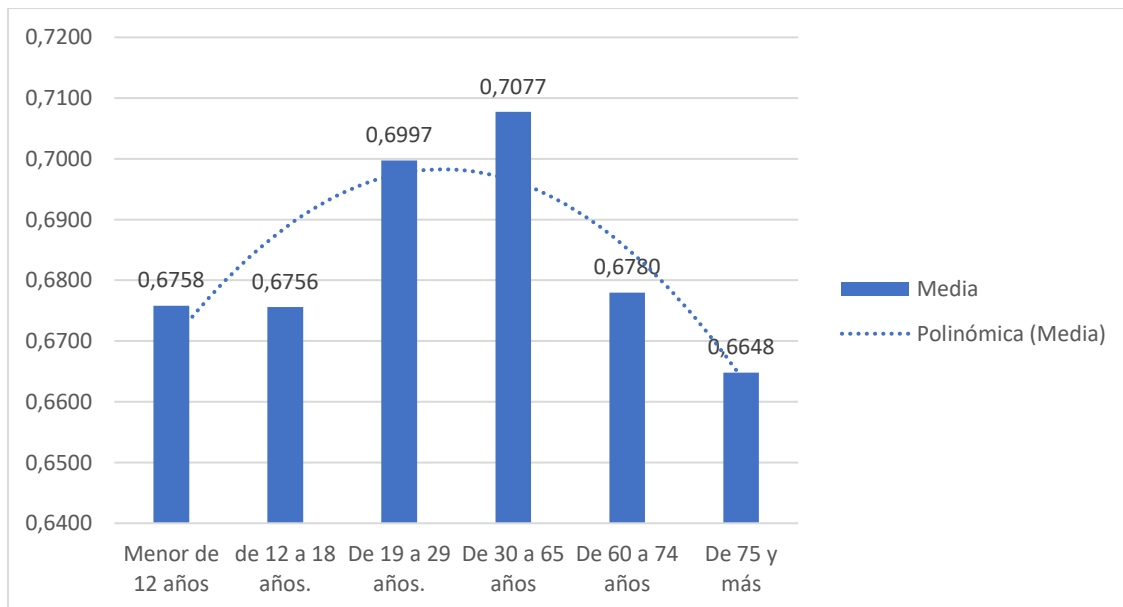
Gráfico 21. Índice promedio RUI por edades simples- Bogotá-2021



Fuente: Procesamiento propio de la EMB 2021.

Las desigualdades en los ingresos por grupos de edad también fueron analizadas anteriormente a partir de la información individual, mostrando apreciables desigualdades en los ingresos por esta variable. Esta tendencia se constata, de una manera más clara, al observar los promedios del índice por grupos de edad (Gráfico 22).

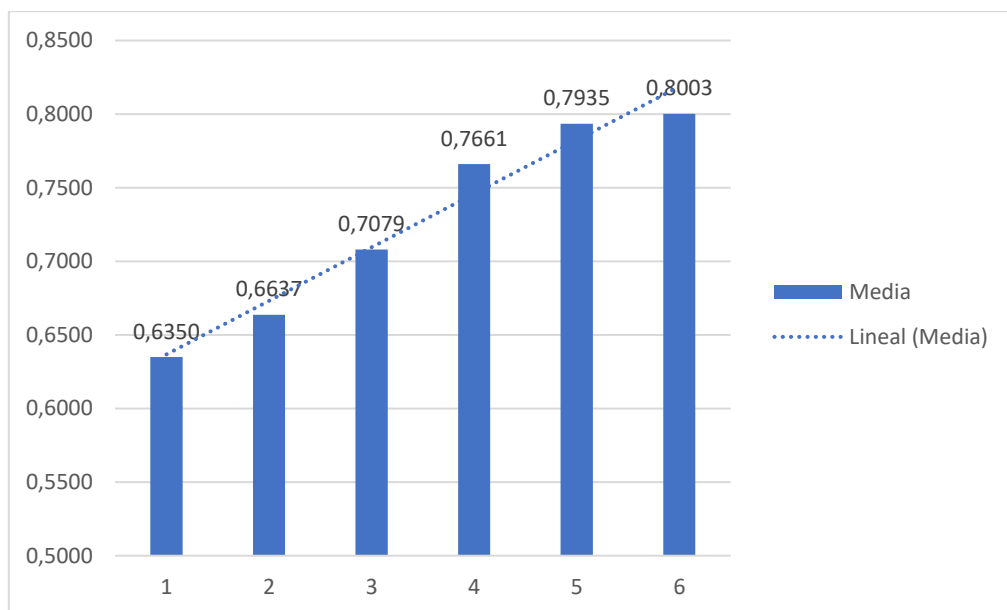
Gráfico 22. Índice RUI promedio por grupos de edad- Bogotá-2021



Fuente: Procesamiento propio de la EMB 2021.

Los valores del índice están igualmente correlacionados directamente con el estrato socioeconómico para el cobro de las tarifas de los servicios públicos domiciliarios. A medida que aumenta el estrato, como tendencia, se incrementa el valor promedio del índice (Gráfico 23).

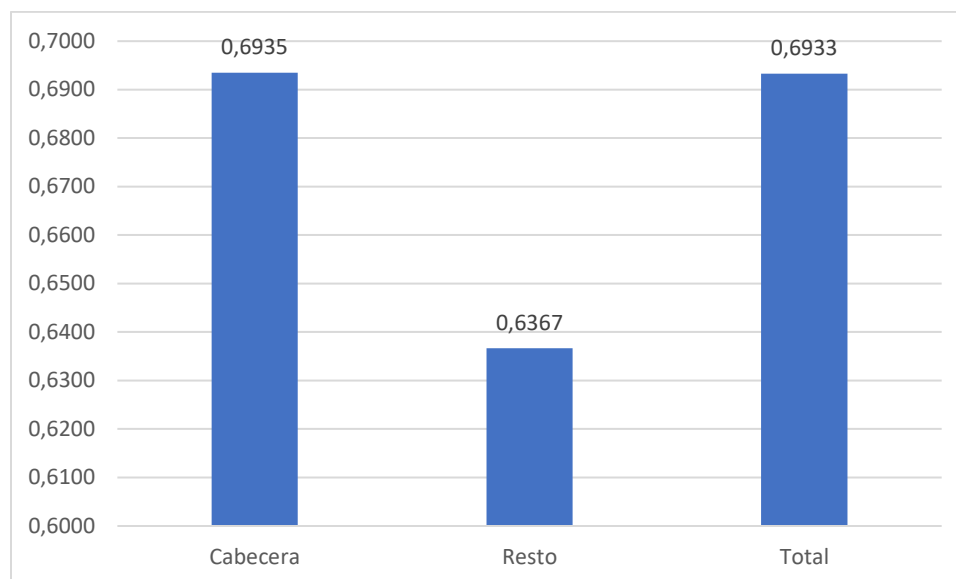
*Gráfico 23. Índice promedio RUI por estratos socioeconómicos vigentes-Bogotá- 2021.*



*Fuente: Procesamiento propio de la EMB 2021.*

De una forma semejante, se evidencian las brechas urbano-rurales en los valores del índice entre estos dos ámbitos. El índice promedio de las habitantes de las zonas rurales es de 0.6367 y el de los de las zonas urbanas 0.6933 (cabecera municipal y resto; Gráfico 24).

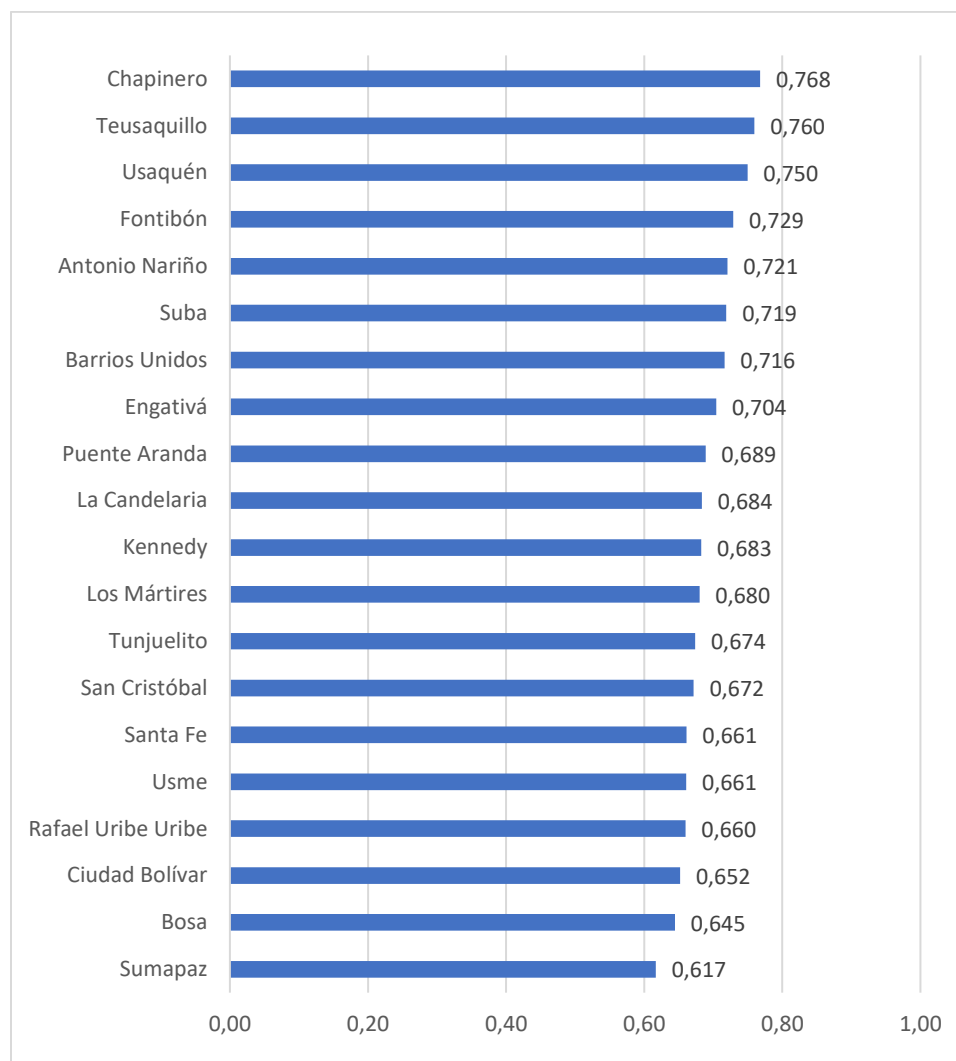
*Gráfico 24. Índice promedio RUI por zonas-Bogotá- 2021.*



*Fuente: Procesamiento propio de la EMB 2021.*

El cálculo del índice promedio RUI por localidades es coincidente con lo que se encuentra en indicadores como el de incidencia de pobreza monetaria o los relativos a calidad de vida. Las localidades con más desventajas en sus condiciones de vida son las mismas que tienen niveles inferiores en el índice. Dentro de ellas se encuentran las localidades ubicadas en el sur de la ciudad: Bosa, Usme, Ciudad Bolívar. Y las localidades con mejores condiciones se encuentran en el norte y centro: Usaquén, Chapinero y Teusaquillo (Gráfico 25).

*Gráfico 25. Índice RUI promedio por localidades- Bogotá-2021*



*Fuente: Procesamiento propio de la EMB 2021.*

## 2.2 Semaforización

Se presenta a continuación la semaforización para el modelo 1 de segmentación, que es uno de los recomendados junto con el 2, con los valores del índice en términos de escala entre cero y uno, y en sus límites superior e inferior, en términos absolutos (Cuadro 1). Para ilustrar una hipotética asignación de subsidios e imposición de contribuciones, se asume que los grupos de los pobres 0, 1 y 2 reciben subsidios, los grupos 3 y 4 no reciben subsidios ni aportan contribuciones, y los grupos 5 y 6 aportan contribuciones.

*Cuadro 1. Semaforización por Grupos de clasificación según IRUI y Nivel de ingresos para el Modelo propuesto 1*

Grupo	IRUI	Nivel de Ingreso	Porcentaje de población	
G1_0	0	0	12,6%	Con subsidio
G1_1	0, 01 a 0,63	0,1 a 203.571		
G1_2	0,6301 a 0,69	203.571 a 483.000	22,7%	
G1_3	0,6901 a 0,74	483.001 a 1.004.313	28,1%	Sin subsidio, ni contribución
G1_4	0,7401 a 0,78	1.004.314 a 2.159.375	20,2%	
G1_5	0,7801 a 0,83	2.159.375 a 5.062.917	12,2%	Con contribución
G1_6	Mayor a 0,8301	Mayor a 5.062.917	4,2%	

*Fuente: Procesamiento propio de la EMB 2021.*

### 3. CARACTERIZACIÓN DE LOS GRUPOS DEL ÍNDICE DE INGRESOS

Se lleva a cabo, en primera instancia, una caracterización de los grupos de uno de los modelos recomendados de la segmentación del índice de ingresos corrientes secundarios (modelo 2), con algunas variables e indicadores económicos y sociodemográficos. Se busca establecer con este procedimiento en qué forma la clasificación hecha con los ingresos diferencia grupos que se distinguen no solo por sus niveles de ingreso sino también por algunos atributos que los identifican.

Como punto de partida, se exploran variables que muestran las diferencias entre los grupos en el campo demográfico. Luego se tratan indicadores que expresan condiciones de vulnerabilidad: discapacitados, en el campo de la vivienda y la falta de afiliación a la seguridad social en salud. A continuación, se adelanta un ejercicio comparativo entre los grupos de los cinco modelos con las variables del Índice de Pobreza Multidimensional. Por último, se abordan las diferencias en la participación de la población en los grupos del índice, en sus expresiones territoriales, por UPZ.

Se presenta posteriormente una comparación de los grupos segmentados en los cinco modelos tomando en cuenta el Índice de Pobreza multidimensional y los 15 indicadores que constituyen sus componentes. Se cubren así dimensiones diversas de las condiciones socioeconómicas de estos grupos.

Por último, se examina la composición de los grupos de los cinco modelos en las UPZ del Distrito Capital para captar su diversidad espacial y algunas particularidades de la aplicación de estos modelos.

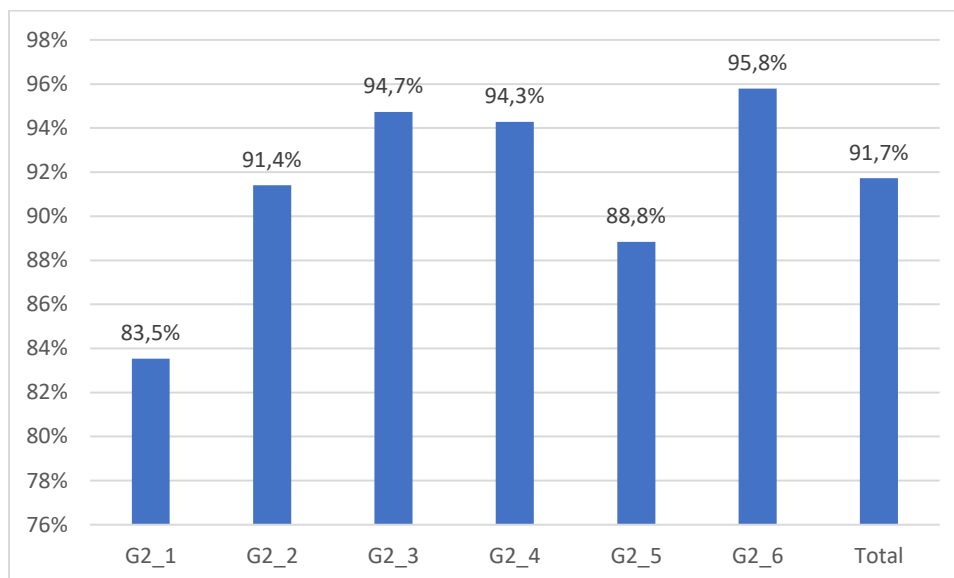
#### 3.1. Características demográficas

Se perciben rasgos distintivos de los grupos del modelo 2 en su composición demográfica. La ubicación dentro de las escalas de ingreso, expresadas en la segmentación de los grupos, están asociadas con disparidades poblacionales.

En cuanto a la composición por sexo de los grupos, la razón de masculinidad, que muestra el número de hombres que hay por cada 100 mujeres, se encuentran peculiaridades entre los grupos. En el inferior y en el quinto las razones son más bajas (menores al 90%), siendo más amplias en los intermedios y el superior. Los valores mayores corresponden con los ingresos más elevados que perciben los hombres. No obstante, en este resultado inciden otros factores relacionados también con características demográficas (Gráfico 26).

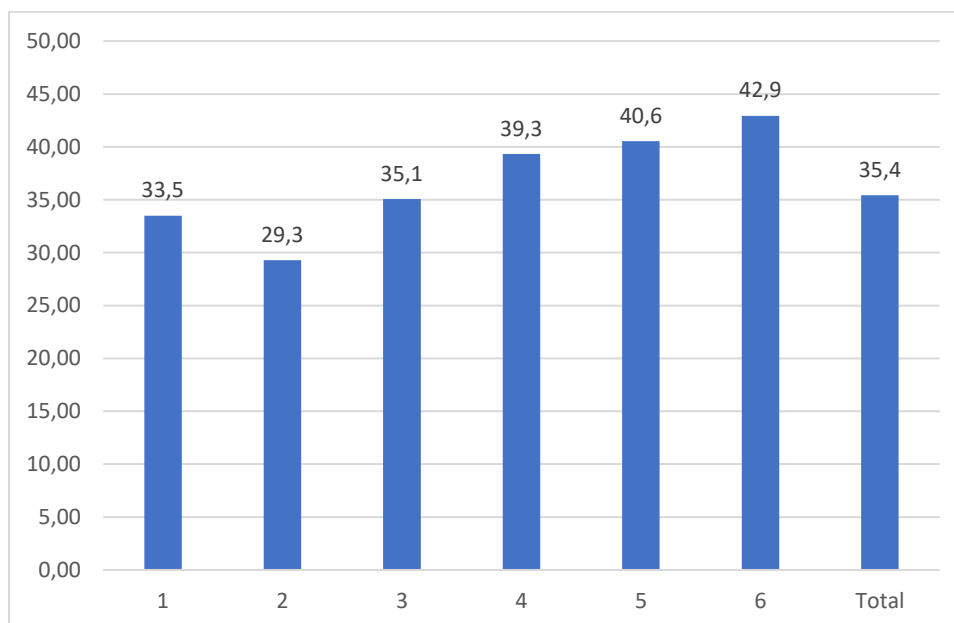
Uno de ellos es el de la composición etaria de los grupos, que también está asociada con desigualdades entre los ingresos, como se ha señalado anteriormente. Se pone de presente en el promedio de edad de los grupos, que tiende a crecer con el aumento de ingreso entre ellos (Gráfico 27). En el grupo 1 la media de la edad es de 33.5 años, en el 2 de 29.3, y en el quinto de 42.9. En tanto aumenta el grupo, la composición de la población es más vieja.

Gráfico 26. Razón de masculinidad, por grupos del modelo 2 del índice RUI -Bogotá-2021



Fuente: Procesamiento propio de la EMB 2021.

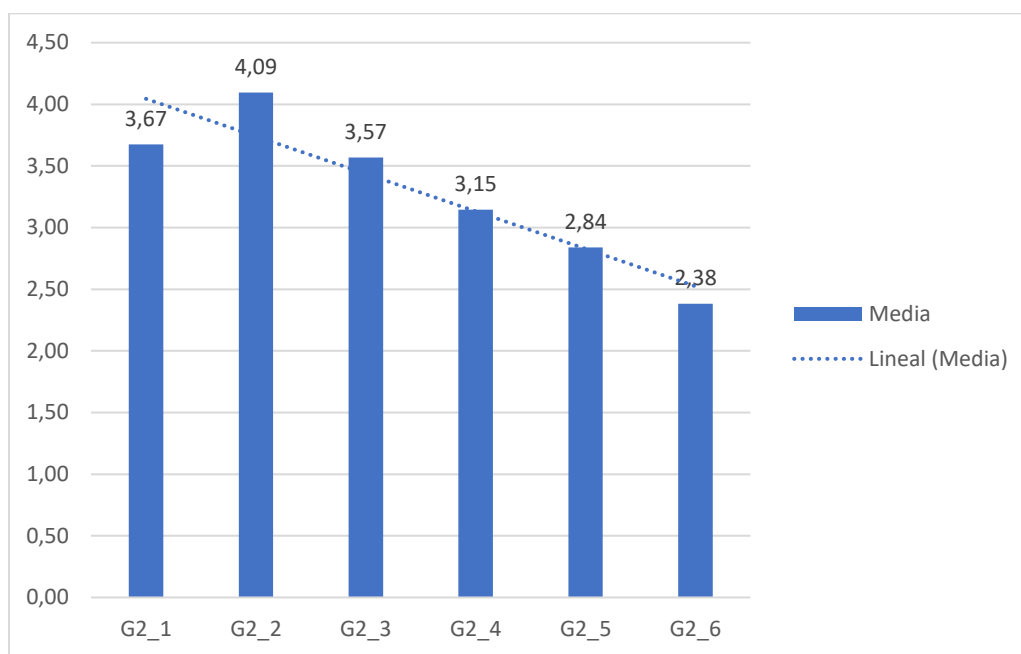
Gráfico 27. Edad promedio, por grupos del modelo 2 del índice RUI -Bogotá-2021



Fuente: Procesamiento propio de la EMB 2021.

El otro factor es el tamaño del hogar que está igualmente relacionado con los niveles de ingreso. Los individuos que se ubican en los grupos con ingresos per cápita inferiores, según el modelo 2, pertenecen, en general, a hogares más numerosos. Los hogares del grupo 2 tienen, en promedio, 4,1 miembros, mientras que los del grupo 5 tiene en promedio 2,4 miembros (Gráfico 28).

*Gráfico 28. Promedio del número de personas por hogar, por grupos del modelo 2 del índice RUI -Bogotá-2021*

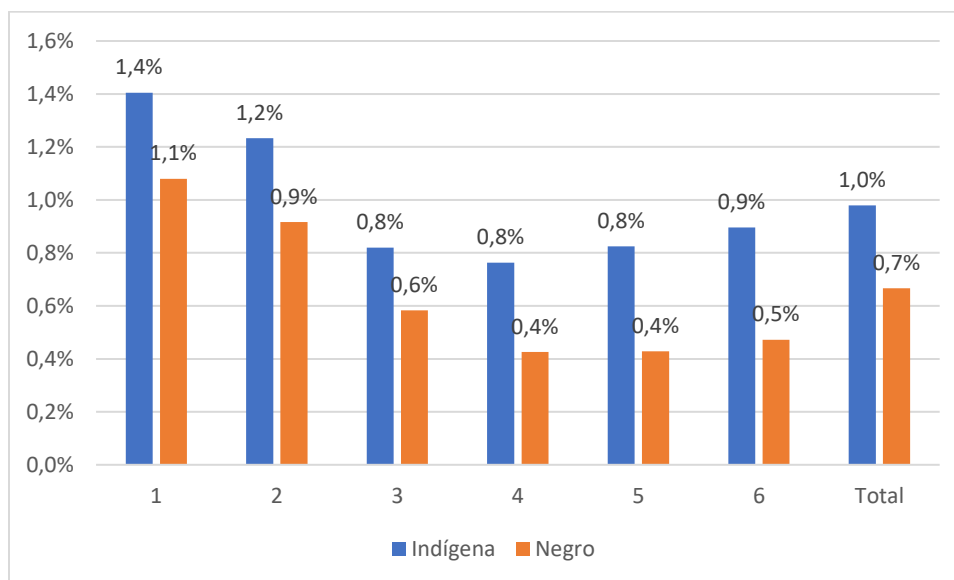


*Fuente: Procesamiento propio de la EMB 2021.*

De otra parte, se encuentran asimismo participaciones diferenciales de las etnias minoritarias en los grupos del modelo 2. En los dos primeros grupos, el porcentaje de población indígena y negra, que no llega al 2% en el conjunto de Bogotá, exhibe valores más altos (Gráfico 29).



Gráfico 29. Porcentaje de población indígena y negra, por grupos del modelo 2 del índice RUI -Bogotá-2021



Fuente: Procesamiento propio de la EMB 2021.

### 3.2. Condiciones de vulnerabilidad

Se analizan algunos indicadores de vulnerabilidad en los grupos del modelo 2 en una doble perspectiva. En primer lugar, para poner de presente las diferencias entre estos grupos en algunos campos relevantes como el de la salud, la vivienda y las condiciones demográficas de presencia de niños y adultos mayores en los hogares. La perspectiva en este caso es la de caracterizar los grupos y poner de presente las diferencias entre ellos. En segunda instancia, se busca ilustrar la forma como podría aplicarse la focalización que se ha calificado de *mixta*<sup>16</sup>, combinando la ubicación en los grupos de ingresos, con indicadores de vulnerabilidad y privación que afectan a los hogares. La intención de este ejercicio no es definir un criterio de focalización particular sino, a través de un ejemplo, mostrar una posible forma de relacionar la información sobre los ingresos y la de carácter sociodemográfico para los procesos de focalización.

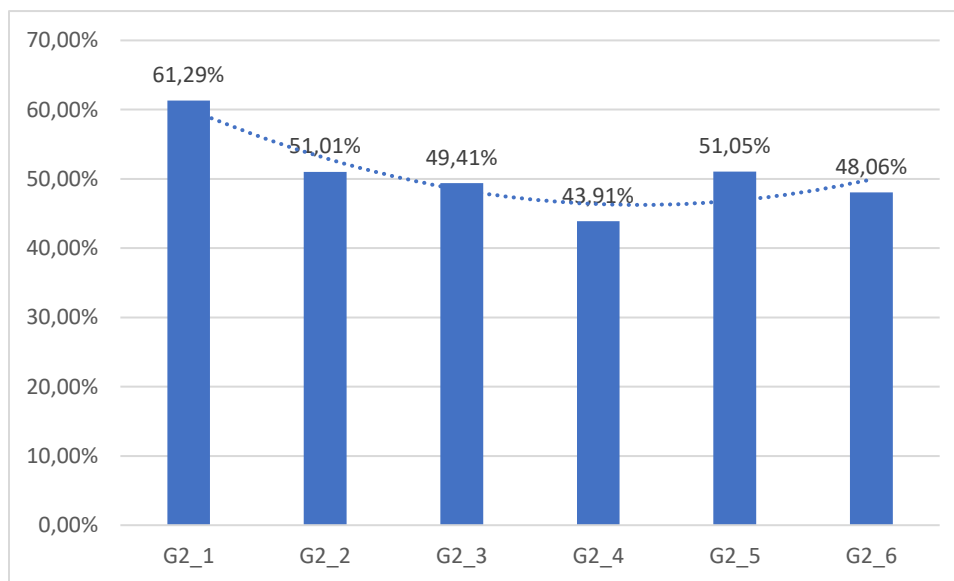
En los gráficos colocados a continuación se exponen los valores que adoptan cinco indicadores de vulnerabilidad que se privilegian, dentro de los grupos por niveles de ingresos del modelo 2 (Gráficos 30 a 34).

<sup>16</sup> Véase el numeral 5.3.1. (Ampliación del índice con información socioeconómica sobre vulnerabilidad y privación) del documento del producto 3 del proyecto: *Diseño conceptual y metodológico para la construcción del Registro Universal de Ingresos*.

Los dos primeros gráficos (30 y 31) se refieren a la presencia de menores de edad y personas con más de 60 años, sin ingresos. Ambos muestran cómo en los grupos de ingresos per cápita inferiores hay mayor proporción de niños y adultos mayores, que disminuye en los grupos intermedios (2 a 4) y aumenta en los superiores (5 y 6). Este comportamiento puede interpretarse en relación con las fases del ciclo de vida familiar y el tamaño de los hogares. En las iniciales, la presencia de niños es frecuente y corresponde con hogares numerosos, con familias extensas donde hay presencia de varias generaciones, incluyendo, con alguna regularidad, mayores e ingresos por persona reducidos. En las etapas posteriores disminuyen el número de niños predominan las familias nucleares y el ingreso asciende, al mismo tiempo que aumenta la proporción de mayores.

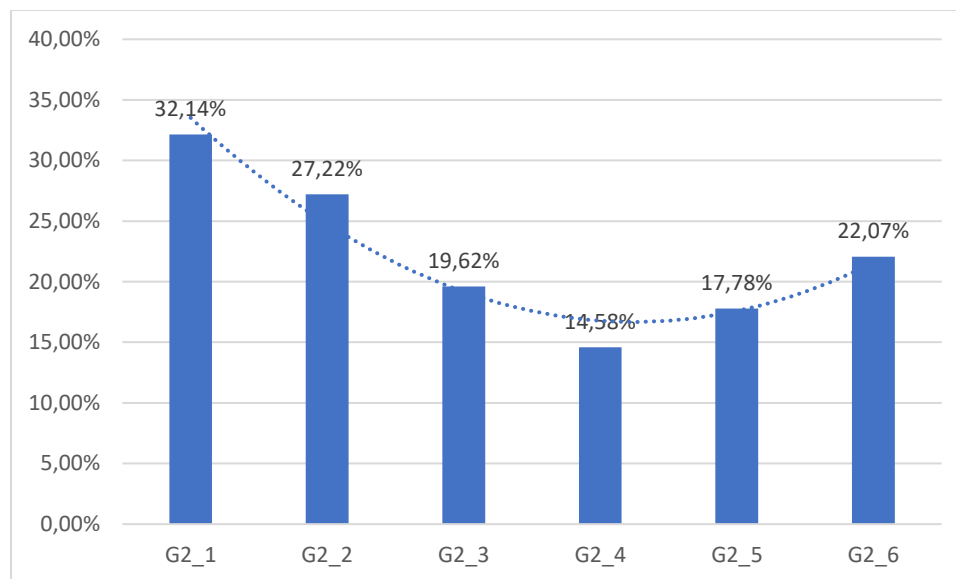
La combinación de un criterio de focalización, ligando los ingresos y la presencia de niños y mayores, permite precisar los grupos poblacionales pertinentes en los segmentos de ingresos que requieren de ayudas por parte de ciertos programas sociales públicos. En algunos de ellos contar con ese tipo de información puede ser imprescindible para su diseño, ejecución y evaluación.

*Gráfico 30. Porcentaje de personas en hogares con menores de edad por pertenencia a los grupos del índice de ingreso del Modelo 2- Bogotá - 2021*



*Fuente: Procesamiento propio de la EMB 2021.*

Gráfico 31. Porcentaje de personas en hogares con mayores de 60 años sin ingreso por pertenencia a los grupos del índice de ingreso del Modelo 2- Bogotá - 2021

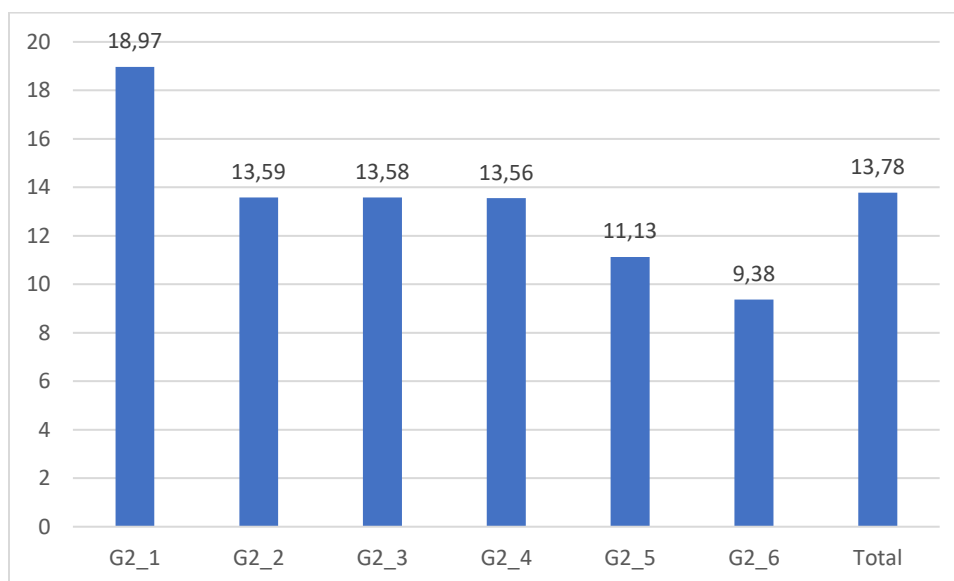


Fuente: Procesamiento propio de la EMB 2021.

Pasando al campo de las condiciones de salud, se observa que, exceptuando el primer grupo y el último, los hogares con alguno de sus miembros discapacitado disminuye a medida que se hacen más altos los ingresos (Gráfico 32).

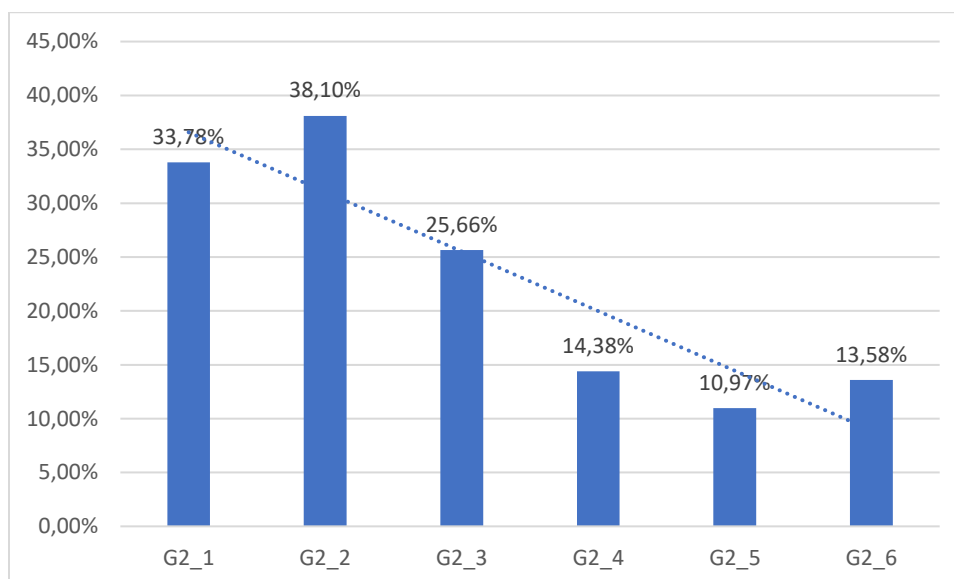
La elevada incidencia de esta población en los grupos con menores ingresos (34% y 38%) corresponde con programas que la privilegian y que necesitarían poder identificarla desde el RUI. Se muestra una tendencia a reducirse las tasas a medida que los ingresos se elevan a pesar de que las edades promedio aumentan con el grupo de ingreso (Gráfico 33).

Gráfico 32. Porcentaje de población con discapacidad por grupos del modelo 2 del índice RUI -Bogotá-2021



Fuente: Procesamiento propio de la EMB 2021.

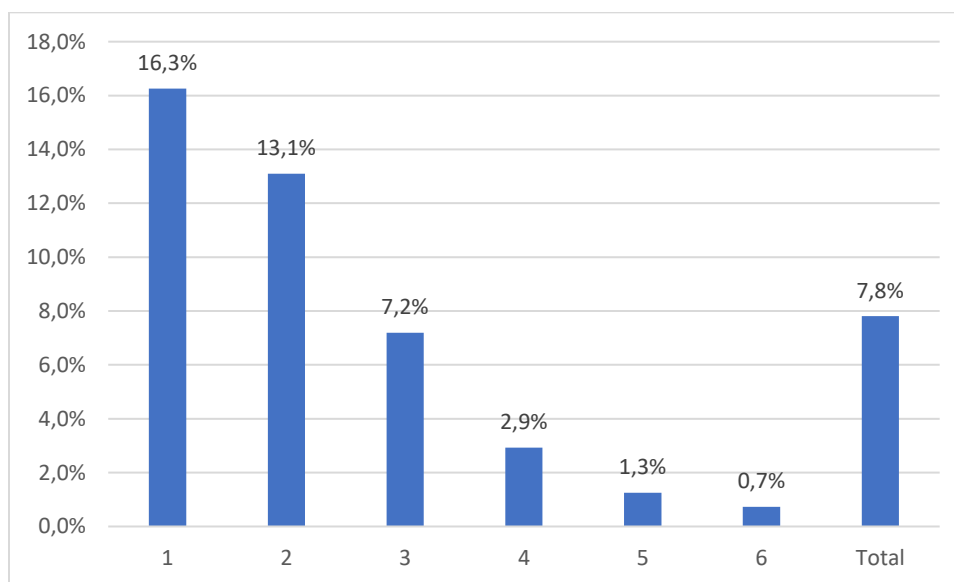
Gráfico 33. Porcentaje de personas en hogares con discapacidad por pertenencia a los grupos del índice de ingreso del Modelo 2- Bogotá – 2021



Fuente: Procesamiento propio de la EMB 2021.

El Gráfico 34, por su parte, muestra los porcentajes de personas sin afiliación a la seguridad social en salud. También en este ámbito se comprueba que son los grupos con menores ingresos los que tiene mayor proporción de no afiliados. Se señala igualmente un segmento vulnerable que requiere de tratamiento especial.

*Gráfico 34. Porcentaje de población sin afiliación a la Seguridad Social en Salud, por grupos del modelo 2 del índice RUI -Bogotá-2021*

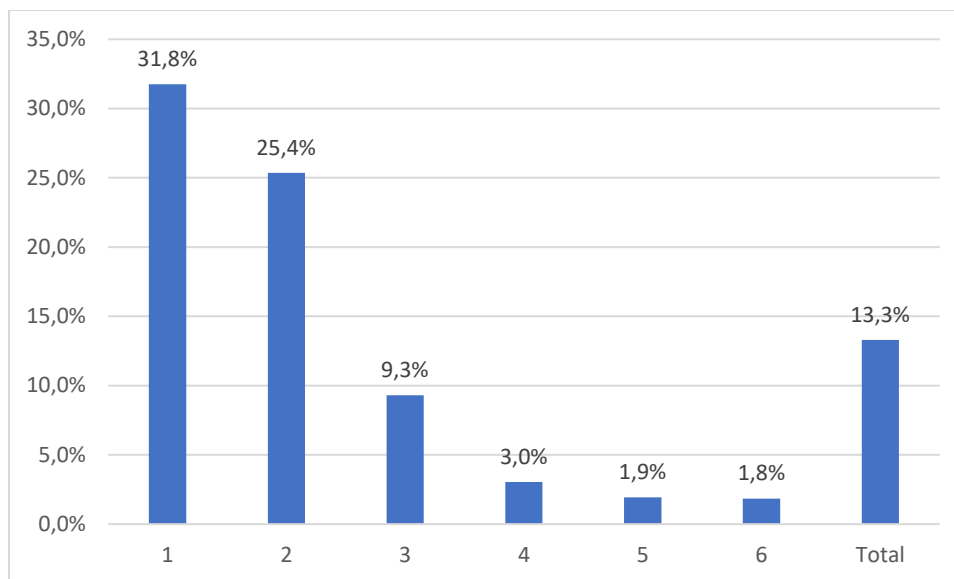


*Fuente: Procesamiento propio de la EMB 2021.*

Se observa, desde otra perspectiva, un patrón claro de disminución de la proporción de hogares con signos de posible desventaja en sus condiciones habitacionales al no disponer de una vivienda propia con el aumento de los ingresos (Gráfico 35).

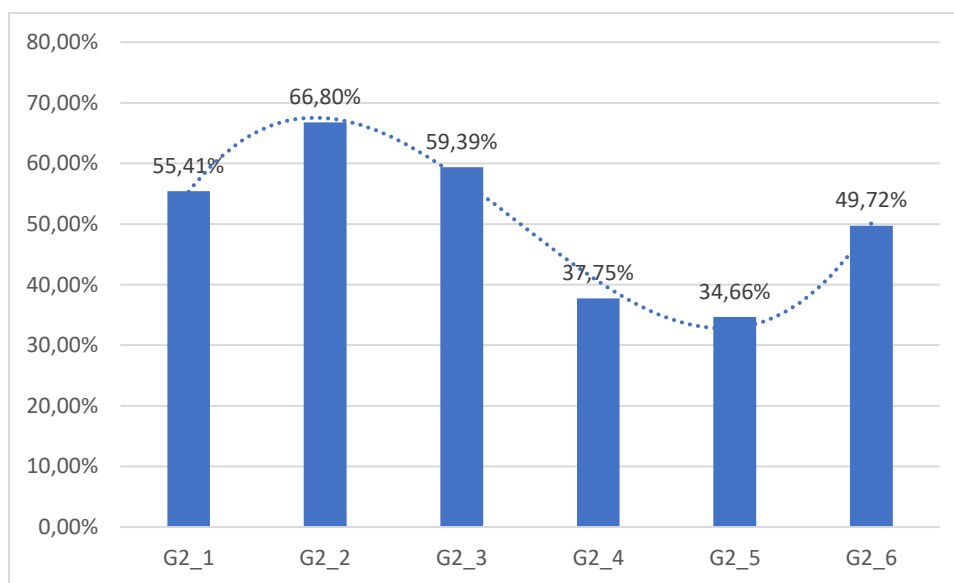
Se distinguen dos situaciones. Una que diferencia a sectores pobres o desfavorecidos que no viven en vivienda propia: 55% en el primer grupo, 67% en el segundo y 59% en el tercero. Otra que abarca a los sectores medios y alto donde no necesariamente esta situación es expresión de una situación desfavorable, sino que puede ser parte de una estrategia que favorezca sus intereses (Gráfico 36).

**Gráfico 35. Porcentaje de población en viviendas con déficit habitacional, por grupos del modelo 2 del índice RUI -Bogotá-2021**



*Fuente: Procesamiento propio de la EMB 2021.*

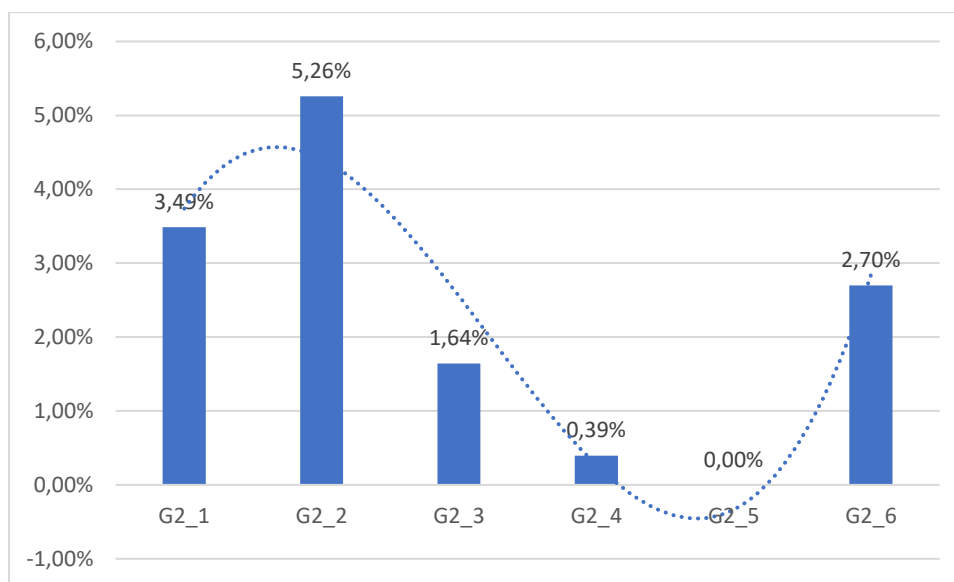
**Gráfico 36. Porcentaje de personas en hogares con hacinamiento por pertenencia a los grupos del índice de ingreso del Modelo 2- Bogotá - 2021**



*Fuente: Procesamiento propio de la EMB 2021.*

Como expresión de las condiciones habitacionales, el porcentaje de personas en hacinamiento que tiene magnitudes menores a los otros indicadores asciende del primer al segundo grupo, disminuye hasta el quinto y exhibe un aumento en el sexto. No obstante, la medida de déficit habitacional, que da cuenta de una manera más global de la calidad de la vivienda, disminuye con el orden del grupo de acuerdo con sus niveles en el índice de ingreso (Gráfico 37). La consideración de estos indicadores contribuye a definir para casos particulares de programas sociales el criterio de focalización.

*Gráfico 37. Porcentaje de personas en hogares con hacinamiento por pertenencia a los grupos del índice de ingreso del Modelo 2- Bogotá - 2021*



*Fuente: Procesamiento propio de la EMB 2021.*

Estos resultados sustentan que los grupos obtenidos con la segmentación propuesta con el ingreso corriente secundario discriminan adecuadamente a agrupaciones poblacionales con niveles diferenciados de acuerdo con sus condiciones de vida, asociadas a expresiones de vulnerabilidad.

Una expresión de la intensidad de la vulnerabilidad es la cantidad de situaciones que afectan simultáneamente a los hogares. Al obtener ese conteo, por grupos de los niveles de ingreso del modelo 2 pueden diferenciarse segmentos que requerirían de intervenciones especiales sobre las condiciones de vida que dependen de los ingresos, y sobre la presencia de factores que consideran la disponibilidad de otros recursos y las estrategias que adoptan las familias para hacer frente a los impactos que las afectan (Pizarro, 2001). En el Cuadro 2 se presenta la frecuencia del número de ocurrencias de las cinco situaciones de vulnerabilidad analizadas en los gráficos 30 a 34 por hogar. Se observa cómo en los tramos de ingresos bajos los porcentajes de acumulación en 3, 4 y 5 eventos es más alto. En el grupo 1 se concentran en esas casillas el 22% de los hogares, en grupo 2 el 25% y

en el 3 el 14%, mientras que en los grupos superiores son del 4.9% en el grupo 4, 3.2% en el 5 y 5.3% en el 6. Información como esta permite adoptar criterios de focalización más acordes con las características y objetivos de los distintos programas sociales.

*Cuadro 2. Porcentaje de personas por número de situaciones de vulnerabilidad que afectan sus hogares por grupos de ingreso del modelo 2- Bogotá- 2021.*

Grupos modelo 2	0	1	2	3	4	5
G2_1	5,64%	28,25%	44,14%	18,45%	3,38%	0,15%
G2_2	6,03%	29,20%	39,95%	20,20%	4,41%	0,21%
G2_3	11,07%	37,90%	37,14%	12,09%	1,74%	0,06%
G2_4	21,43%	51,15%	22,54%	4,72%	0,16%	0,00%
G2_5	19,71%	50,50%	26,65%	1,93%	1,22%	0,00%
G2_6	17,32%	34,48%	42,95%	5,25%	0,00%	0,00%

*Fuente: Procesamiento propio de la EMB 2021.*

### 3.3. Índice de pobreza multidimensional (IPM) y sus componentes por grupos poblacionales de los modelos 1 a 5 del índice de ingresos.

El Índice de Pobreza Multidimensional (IPM) y los 15 indicadores que conforman sus componentes, contribuyen a completar la caracterización de los grupos del modelo 2, que se ha tratado en las secciones anteriores, y, en una perspectiva comparativa, los de los otros cuatro modelos de agrupación (Cuadros 3 a 7).

Los indicadores de cada componente expresan el porcentaje de persona en cuyos hogares se encuentran situaciones de privación, que son tenidas en cuenta para obtener la medida de la pobreza multidimensional tomando en cuenta distintas dimensiones socioeconómicas.

En general, se encuentra que, como en las temáticas tratadas anteriormente, hay una asociación entre los niveles de ingreso propios de los grupos poblacionales de los distintos modelos y la magnitud de los indicadores examinados. Con pocas excepciones, los indicadores que se refieren a condiciones de carencia disminuyen con los ingresos en los campos tratados por el IPM: condiciones educativas del hogar, condiciones de la niñez y la juventud, trabajo, salud, y servicios públicos domiciliarios y condiciones de la vivienda. Esto sucede con los grupos de todos los modelos, sin que esta regla se vea afectada por la forma de agrupamiento de la población.

Dentro de las irregularidades frente a este comportamiento se encuentra que, con alguna frecuencia, el primer grupo presenta condiciones más favorables que el segundo. Esta situación señala que el primer grupo de los modelos, sin ingreso o con ingresos bajos, tiene posiblemente una composición donde coexisten pobres “estructurales”, con pobres



“coyunturales” que conservan algunas de sus condiciones favorables de vida anteriores cuando sufren una caída temporal de sus ingresos.

El Índice global de pobreza multidimensional que recoge el aporte de los indicadores, resulta una expresión sintética de ellos y da cuenta de la ubicación de los grupos dentro de una escala multidimensional. En cada modelo varían las brechas entre los grupos. En los cuatro primeros, no hay diferencias notables en los resultados que permitan adoptar un juicio para establecer cuál modelo es más adecuado. En cuanto al quinto modelo, sí es notable que tiene menor capacidad para distinguir a los grupos intermedios (Cuadros 3 a 7).

**Cuadro 3. Indicadores de los componentes del Índice de Pobreza Multidimensional-Modelo 1 del Índice de Ingresos- Bogotá- 2021**

Modelo	Grupo dentro del modelo	Bajo logro educativo (Componente de IPM)	Analfabetismo (Componente de IPM)	Inasistencia escolar (Componente de IPM)	Rezago escolar (Componente de IPM)	Atención a la primera infancia (Componente de IPM)	Trabajo infantil (Componente de IPM)	Desempleo de larga duración (Componente de IPM)	Empleo informal (Componente de IPM)	Aseguramiento a salud (Componente de IPM)	Acceso a servicios de salud (Componente de IPM)	Acceso a servicio de agua (Componente de IPM)	Acceso a servicio de alcantarillado (Componente de IPM)	Condición de los pisos de la vivienda (Componente de IPM)	Condición de paredes de la vivienda (Componente de IPM)	Condición de hacinamiento de la vivienda (Componente de IPM)	Índice de Pobreza Multidimensional (Población en condición de pobreza multidimensional)
Modelo 1 Ingreso secundario	0	38,3%	4,8%	3,1%	22,5%	6,7%	0,2%	66,3%	95,1%	23,1%	1,4%	0,5%	0,9%	0,5%	0,9%	9,3%	18,2%
	1	44,3%	6,5%	7,1%	42,9%	18,8%	0,9%	18,0%	92,6%	31,1%	4,4%	1,5%	2,4%	1,6%	2,1%	24,7%	21,7%
	2	30,9%	3,6%	4,5%	36,6%	15,1%	1,0%	10,8%	71,2%	25,1%	3,5%	0,4%	0,8%	0,4%	0,4%	13,8%	9,6%
	3	19,8%	2,4%	1,4%	21,6%	8,0%	0,8%	9,5%	60,7%	15,9%	2,5%	0,3%	0,4%	0,1%	0,2%	4,3%	2,5%
	4	10,1%	1,3%	0,6%	14,4%	5,6%	0,2%	10,6%	42,3%	7,4%	2,0%	0,0%	0,1%	0,0%	0,0%	1,5%	0,6%
	5	2,6%	0,5%	0,3%	13,8%	5,6%	0,1%	10,5%	27,0%	3,4%	1,2%	0,0%	0,3%	0,0%	0,1%	1,1%	0,1%
	6	0,8%	0,4%	0,1%	10,2%	4,4%	0,0%	12,0%	17,5%	2,1%	0,7%	0,0%	0,0%	0,0%	0,1%	0,8%	0,0%
	Total	20,3%	2,6%	2,3%	23,8%	9,5%	0,6%	13,7%	57,6%	15,7%	2,5%	0,3%	0,6%	0,3%	0,4%	7,2%	5,6%

Fuente: Procesamiento propio de la EMB 2021.

**Cuadro 4. Indicadores de los componentes del Índice de Pobreza Multidimensional-Modelo 2 del Índice de Ingresos- Bogotá- 2021**

Modelo	Grupo dentro del modelo	Bajo logro educativo (Componente de IPM)	Analfabetismo (Componente de IPM)	Inasistencia escolar (Componente de IPM)	Rezago escolar (Componente de IPM)	Atención a la primera infancia (Componente de IPM)	Trabajo infantil (Componente de IPM)	Desempleo de larga duración (Componente de IPM)	Empleo informal (Componente de IPM)	Aseguramiento a salud (Componente de IPM)	Acceso a servicios de salud (Componente de IPM)	Acceso a servicio de agua (Componente de IPM)	Acceso a servicio de alcantarillado (Componente de IPM)	Condición de los pisos de la vivienda (Componente de IPM)	Condición de paredes de la vivienda (Componente de IPM)	Condición de hacinamiento de la vivienda (Componente de IPM)	Índice de Pobreza Multidimensional (Población en condición de pobreza multidimensional)
Modelo 2 Ingreso secundario	1	42,00%	5,80%	5,60%	35,10%	14,20%	0,60%	36,50%	93,50%	28,00%	3,30%	1,10%	1,80%	1,20%	1,60%	18,80%	20,40%
	2	30,90%	3,60%	4,50%	36,60%	15,10%	1,00%	10,80%	71,20%	25,10%	3,50%	0,40%	0,80%	0,40%	0,40%	13,80%	9,60%
	3	19,80%	2,40%	1,40%	21,60%	8,00%	0,80%	9,50%	60,70%	15,90%	2,50%	0,30%	0,40%	0,10%	0,20%	4,30%	2,50%
	4	10,10%	1,30%	0,60%	14,40%	5,60%	0,20%	10,60%	42,30%	7,40%	2,00%	0,00%	0,10%	0,00%	0,00%	1,50%	0,60%
	5	2,60%	0,50%	0,30%	13,80%	5,60%	0,10%	10,50%	27,00%	3,40%	1,20%	0,00%	0,30%	0,00%	0,10%	1,10%	0,10%
	6	0,80%	0,40%	0,10%	10,20%	4,40%	0,00%	12,00%	17,50%	2,10%	0,70%	0,00%	0,70%	0,00%	0,10%	0,80%	0,00%
	Total	20,30%	2,60%	2,30%	23,80%	9,50%	0,60%	13,70%	57,60%	15,70%	2,50%	0,30%	0,60%	0,30%	0,40%	7,20%	5,60%

Fuente: Procesamiento propio de la EMB 2021

**Cuadro 5. Indicadores de los componentes del Índice de Pobreza Multidimensional-Modelo 3 del Índice de Ingresos- Bogotá- 2021**

Modelo	Grupo dentro del modelo	Bajo logro educativo (Componente de IPM)	Analfabetismo (Componente de IPM)	Inasistencia escolar (Componente de IPM)	Rezago escolar (Componente de IPM)	Atención a la primera infancia (Componente de IPM)	Trabajo infantil (Componente de IPM)	Desempleo de larga duración (Componente de IPM)	Empleo informal (Componente de IPM)	Aseguramiento a salud (Componente de IPM)	Acceso a servicios de salud (Componente de IPM)	Acceso a servicio de agua (Componente de IPM)	Acceso a servicio de alcantarillado (Componente de IPM)	Condición de los pisos de la vivienda (Componente de IPM)	Condición de paredes de la vivienda (Componente de IPM)	Condición de hacinamiento de la vivienda (Componente de IPM)	Índice de Pobreza Multidimensional (Población en condición de pobreza multidimensional)
Modelo 3 Ingreso secundario	1	38,30%	4,80%	3,10%	22,50%	6,70%	0,20%	66,30%	95,10%	23,10%	1,40%	0,50%	0,90%	0,50%	0,90%	9,30%	18,20%
	2	46,00%	7,20%	7,10%	39,30%	18,20%	0,80%	23,40%	95,40%	30,50%	4,90%	1,70%	2,30%	2,10%	2,40%	24,80%	22,60%
	3	33,10%	4,00%	4,80%	38,40%	16,00%	1,00%	11,30%	74,40%	26,30%	3,60%	0,70%	1,10%	0,60%	0,70%	15,70%	11,70%
	4	18,80%	2,30%	1,40%	20,60%	7,60%	0,70%	9,40%	58,60%	14,90%	2,50%	0,10%	0,30%	0,10%	0,10%	3,90%	2,20%
	5	6,90%	0,90%	0,50%	14,10%	5,50%	0,20%	11,00%	36,60%	5,70%	1,80%	0,00%	0,20%	0,00%	0,10%	1,30%	0,40%
	6	1,20%	0,30%	0,20%	12,00%	5,30%	0,00%	11,10%	20,40%	2,50%	0,80%	0,00%	0,50%	0,00%	0,10%	0,90%	0,10%
	Total	20,30%	2,60%	2,30%	23,80%	9,50%	0,60%	13,70%	57,60%	15,70%	2,50%	0,30%	0,60%	0,30%	0,40%	7,20%	5,60%

Fuente: Procesamiento propio de la EMB 2021

**Cuadro 6. Indicadores de los componentes del Índice de Pobreza Multidimensional-Modelo 4 del Índice de Ingresos- Bogotá- 2021**

Modelo	Grupo dentro del modelo	Bajo logro educativo (Componente de IPM)	Analfabetismo (Componente de IPM)	Inasistencia escolar (Componente de IPM)	Rezago escolar (Componente de IPM)	Atención a la primera infancia (Componente de IPM)	Trabajo infantil (Componente de IPM)	Desempleo de larga duración (Componente de IPM)	Empleo informal (Componente de IPM)	Aseguramiento a salud (Componente de IPM)	Acceso a servicios de salud (Componente de IPM)	Acceso a servicio de agua (Componente de IPM)	Acceso a servicio de alcantarillado (Componente de IPM)	Condición de los pisos de la vivienda (Componente de IPM)	Condición de paredes de la vivienda (Componente de IPM)	Condición de hacinamiento de la vivienda (Componente de IPM)	Índice de Pobreza Multidimensional (Población en condición de pobreza multidimensional)
Modelo 4 Ingreso secundario	1	41,40%	5,60%	5,80%	37,30%	15,60%	0,80%	30,20%	89,60%	28,50%	3,20%	1,10%	1,60%	1,00%	1,40%	18,50%	19,30%
	2	29,20%	3,40%	4,20%	36,30%	14,70%	0,90%	10,50%	69,60%	24,30%	3,60%	0,40%	0,80%	0,40%	0,40%	14,00%	8,60%
	3	22,70%	2,80%	2,10%	25,10%	8,80%	0,70%	9,50%	64,40%	18,80%	2,90%	0,20%	0,40%	0,20%	0,20%	5,10%	3,50%
	4	16,40%	2,10%	0,80%	17,20%	6,90%	0,70%	9,70%	55,50%	12,50%	2,20%	0,10%	0,30%	0,00%	0,10%	3,10%	1,50%
	5	9,50%	1,00%	0,50%	14,30%	5,60%	0,20%	10,90%	41,50%	7,00%	2,00%	0,00%	0,10%	0,00%	0,00%	1,50%	0,50%
	6	2,30%	0,50%	0,30%	12,90%	5,30%	0,10%	10,90%	24,60%	3,10%	1,10%	0,00%	0,40%	0,00%	0,10%	1,00%	0,10%
	Total	20,30%	2,60%	2,30%	23,80%	9,50%	0,60%	13,70%	57,60%	15,70%	2,50%	0,30%	0,60%	0,30%	0,40%	7,20%	5,60%

Fuente: Procesamiento propio de la EMB 2021

**Cuadro 7. Indicadores de los componentes del Índice de Pobreza Multidimensional-Modelo 5 del Índice de Ingresos- Bogotá- 2021**

Modelo	Grupo dentro del modelo	Bajo logro educativo (Componente de IPM)	Analfabetismo (Componente de IPM)	Inasistencia escolar (Componente de IPM)	Rezago escolar (Componente de IPM)	Atención a la primera infancia (Componente de IPM)	Trabajo infantil (Componente de IPM)	Desempleo de larga duración (Componente de IPM)	Empleo informal (Componente de IPM)	Aseguramiento a salud (Componente de IPM)	Acceso a servicios de salud (Componente de IPM)	Acceso a servicio de agua (Componente de IPM)	Acceso a servicio de alcantarillado (Componente de IPM)	Condición de los pisos de la vivienda (Componente de IPM)	Condición de paredes de la vivienda (Componente de IPM)	Condición de hacinamiento de la vivienda (Componente de IPM)	Índice de Pobreza Multidimensional (Población en condición de pobreza multidimensional)
Modelo 5 Ingreso secundario	1	38,10%	5,00%	0,054	37,60%	15,60%	0,80%	23,80%	83,50%	27,70%	3,50%	0,90%	1,40%	0,90%	1,10%	17,80%	16,30%
	2	24,10%	2,80%	0,025	28,10%	10,30%	0,90%	9,90%	65,70%	20,00%	3,00%	0,30%	0,50%	0,20%	0,20%	7,10%	4,60%
	3	14,80%	1,80%	0,008	16,40%	6,40%	0,60%	9,90%	52,20%	10,80%	2,10%	0,10%	0,20%	0,00%	0,10%	2,50%	1,10%
	4	3,90%	0,60%	0,003	13,20%	5,40%	0,10%	11,10%	28,50%	4,20%	1,40%	0,00%	0,30%	0,00%	0,10%	1,10%	0,20%
	Total	20,30%	2,60%	0,023	23,80%	9,50%	0,60%	13,70%	57,60%	15,70%	2,50%	0,30%	0,60%	0,30%	0,40%	7,20%	5,60%

Fuente: Procesamiento propio de la EMB 2021

### 3.4. Distribución por grupos de los modelos 1 a 5 del índice RUI de ingresos por UPZ

A nivel de UPZ se encuentran contrastes destacados en la composición de la población por las agrupaciones propuestas del índice, que se exponen a continuación. En esta sección se desarrolla este tema. La metodología y fuentes de información utilizadas se exponen en el Recuadro 1.

Modelo 1 (Ilustración 1): los grupos económicos con menores ingresos (grupos 0 al 2) se concentran en el sur y suroccidente de la ciudad; en la zona central, occidental y noroccidental predominan los grupos intermedios 3 y 4. Los grupos económicos de más altos ingresos se concentran en la zona norte y este de la ciudad.

Modelo 2 (Ilustración 2): posee una distribución similar a la del modelo 1, con la particularidad de que la presencia de los grupos de bajos ingresos es más notoria y se hace presente en más UPZ que en el modelo anterior. También se destaca que los grupos intermedios pierden presencia en aquellas zonas donde eran más predominantes, concentrándose hacia la parte central y noroeste de la ciudad. Por último, los grupos de más altos ingresos se concentran en el norte y este de la ciudad. Aunque se trate de variables diferentes, visualmente es posible destacar que este modelo es el que más se parece a la distribución de la estratificación vigente de la ciudad.

Modelo 3 (Ilustración 3): representa un aumento de los grupos de ingresos intermedios y, a diferencia de los dos modelos anteriores, la participación de los grupos con más bajos ingresos disminuye en las UPZ donde tenían una mayor representatividad.

Modelo 4 (Ilustración 4): en este modelo se presenta una estabilización en las proporciones de los grupos hacia la zona sur de la ciudad, mientras que los grupos más altos predominan en mayor proporción dentro de las UPZ donde hacen presencia respecto a modelos anteriores.

Modelo 5 (Ilustración 5): posee la particularidad de que tiene menos grupos representando una distribución donde los de más bajos ingresos predominan en el sur, sureste y suroeste de la ciudad, mientras que el grupo de más alto ingreso destaca mucho más que en modelos anteriores en la zona central y norte de la ciudad.

Por último, se añade el mapa de la estratificación actual, por manzanas agrupadas por UPZ, (Ilustración 6) con fines comparativos, de manera visual, con los modelos propuestos dentro de este informe.

### Recuadro 1. Metodología y fuentes de información para el análisis de la distribución de los grupos de los modelos por UPZ

Los productos cartográficos presentados fueron elaborados con los insumos suministrados por la Oficina de Planeación, junto con datos abiertos disponibles en las plataformas digitales estatales, específicamente la base de datos de la Encuesta Multipropósito de 2021 y la capa cartográfica de divisiones de UPZ de la ciudad de Bogotá.

Inicialmente se realiza la asignación de las UPZ a las encuestas de la base de datos con ayuda del identificador de directorio de personas. En este punto se destacan tres observaciones importantes:

- No todos los registros cuentan con su localización UPZ, dando un total de 226.740 registros de los 235.103 con UPZ asignable.
- Llegan códigos UPZ que no se encuentran referenciados en las capas cartográficas, es decir, códigos UPZ que no cuentan con un polígono definido en el espacio actualmente. Esto lleva a que sean 191.358 Los registros que pueden ser mapeados y asignados a un polígono en el espacio.
- UPZ en las que no se encuentran encuestas dentro de la base de datos, las cuales son listadas a continuación:

#### Listado de UPZ sin encuestas dentro de la base de datos multipropósito 2021.

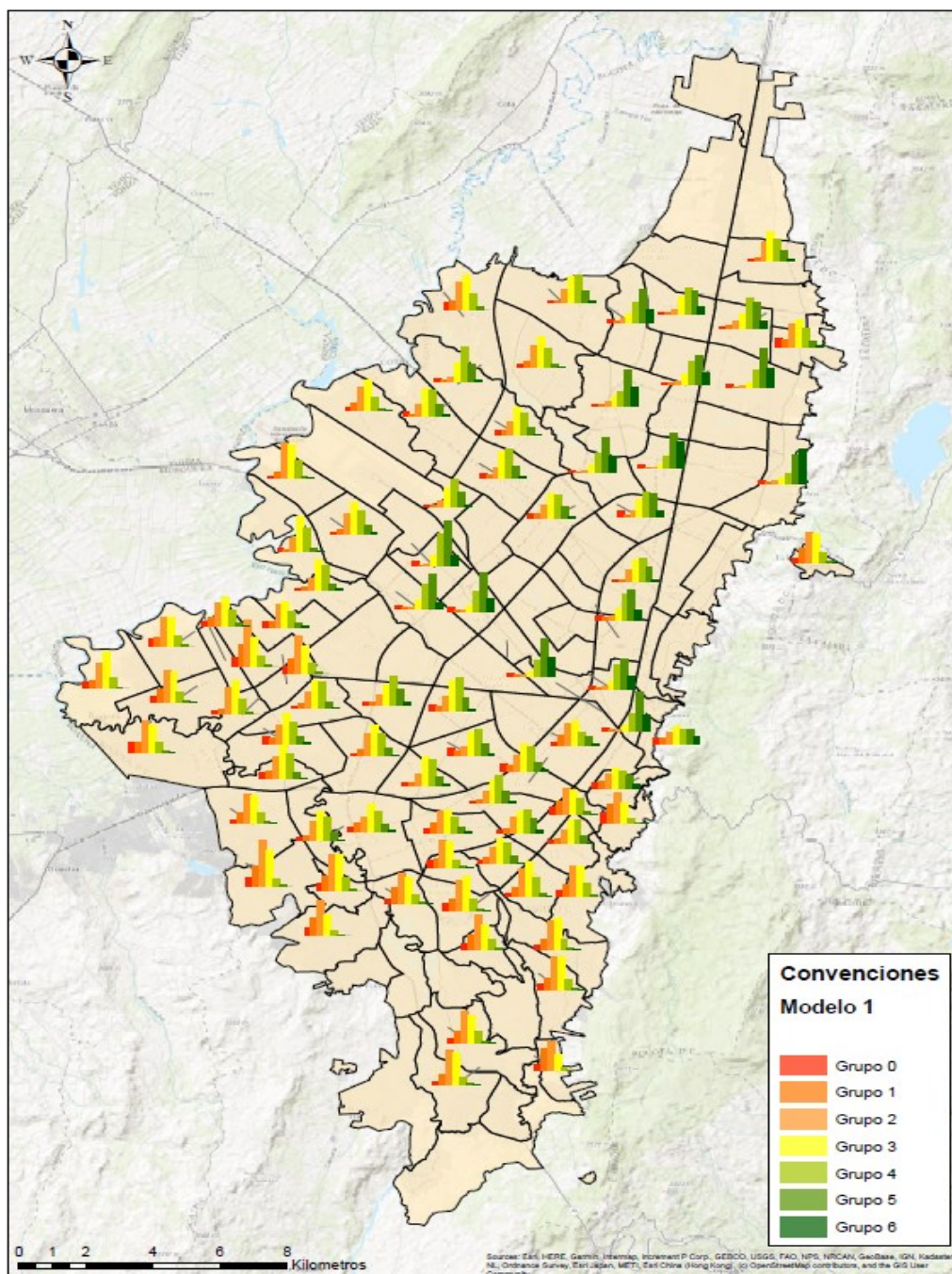
UPZ59	Alfonso López	UPZ56	Danubio
UPZ46	Castilla	UPZ91	Sagrado Corazón
UPZ22	Doce De Octubre	UPZ3	Guaymaral
UPZ108	Zona Industrial	UPZ31	Santa Cecilia
UPZ113	Bavaria	UPZ117	Aeropuerto El Dorado
UPZ99	Chapinero	UPZ88	El Refugio
UPZ10	La Uribe	UPZ104	Parque Simón Bolívar - Can
UPZ61	Ciudad Usme	UPZ60	Parque Entre Nubes
UPZ2	La Academia	UPZ68	El Tesoro
UPZ90	Pardo Rubio	UPZ64	Monte Blanco
UPZ16	Santa Barbara	UPZ106	La Esmeralda
UPZ103	Parque Salitre	UPZ111	Puente Aranda
UPZ93	Las Nieves	UPZ63	El Mochuelo
UPZ1	Paseo De Los Libertadores	UPZ115	Capellanía
UPZ105	Jardín botánico	UPZ15	Country Club
UPZ17	San José De Bavaria	UPZ97	Chico Lago

*Fuente: Elaboración propia.*

Después de asignar el código UPZ a las encuestas de la base de datos es posible asignar la información obtenida de los modelos de grupos económicos a los polígonos de estas, dando como resultado un balance de los diferentes grupos a cada UPZ que puede ser representado a manera de distribución dentro de cada uno de los polígonos que las representan.

Esto da como resultado la generación de 5 mapas (ilustraciones 1 a 5) donde se grafica dicha distribución a lo largo de la ciudad y donde es posible observar zonas donde se concentran en diferentes medidas estos grupos. Adicionalmente se realiza un mapa con la estratificación socioeconómica vigente (ilustración 6), que, a pesar de encontrarse a una escala mucho más definida, puede ser de interés al momento de observar cual de estos modelos se ajusta más a la realidad actual de la ciudad.

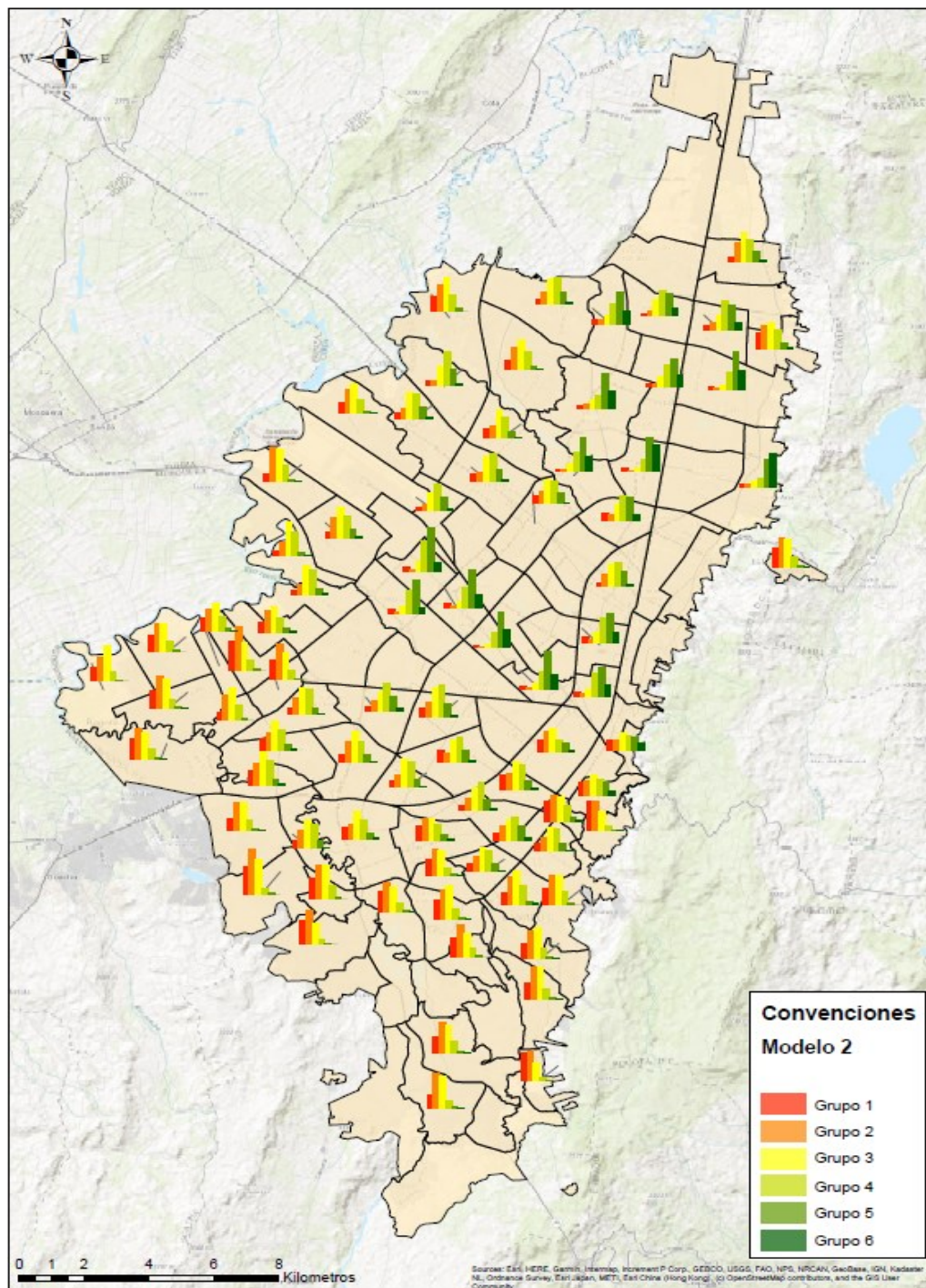
Ilustración 1. Mapa de distribución de grupos del índice de ingreso por UPZ. Modelo 1



Fuente: Elaboración propia.

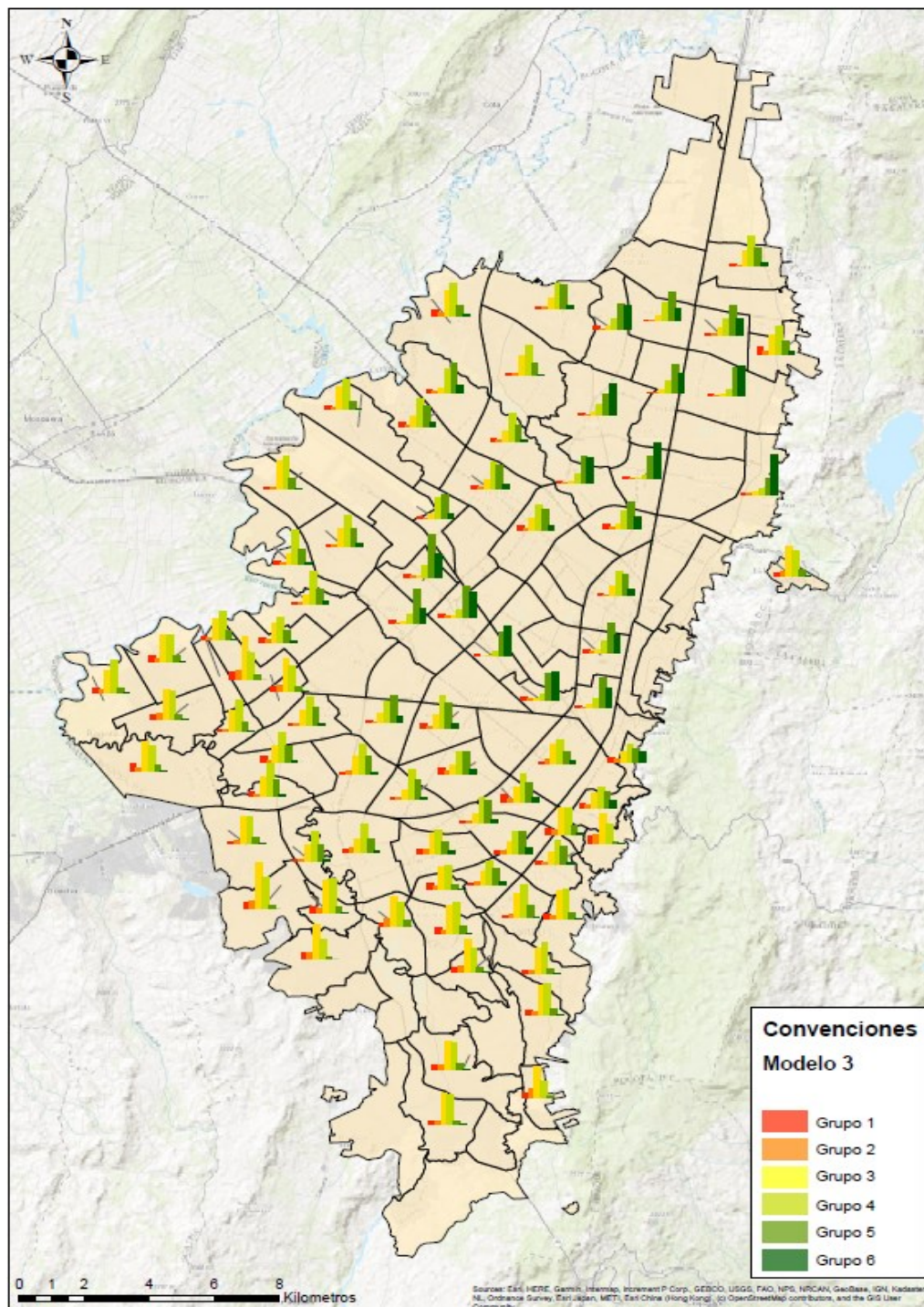


Ilustración 2. Mapa de distribución de grupos del índice de ingreso por UPZ. Modelo 2



Fuente: Elaboración propia.

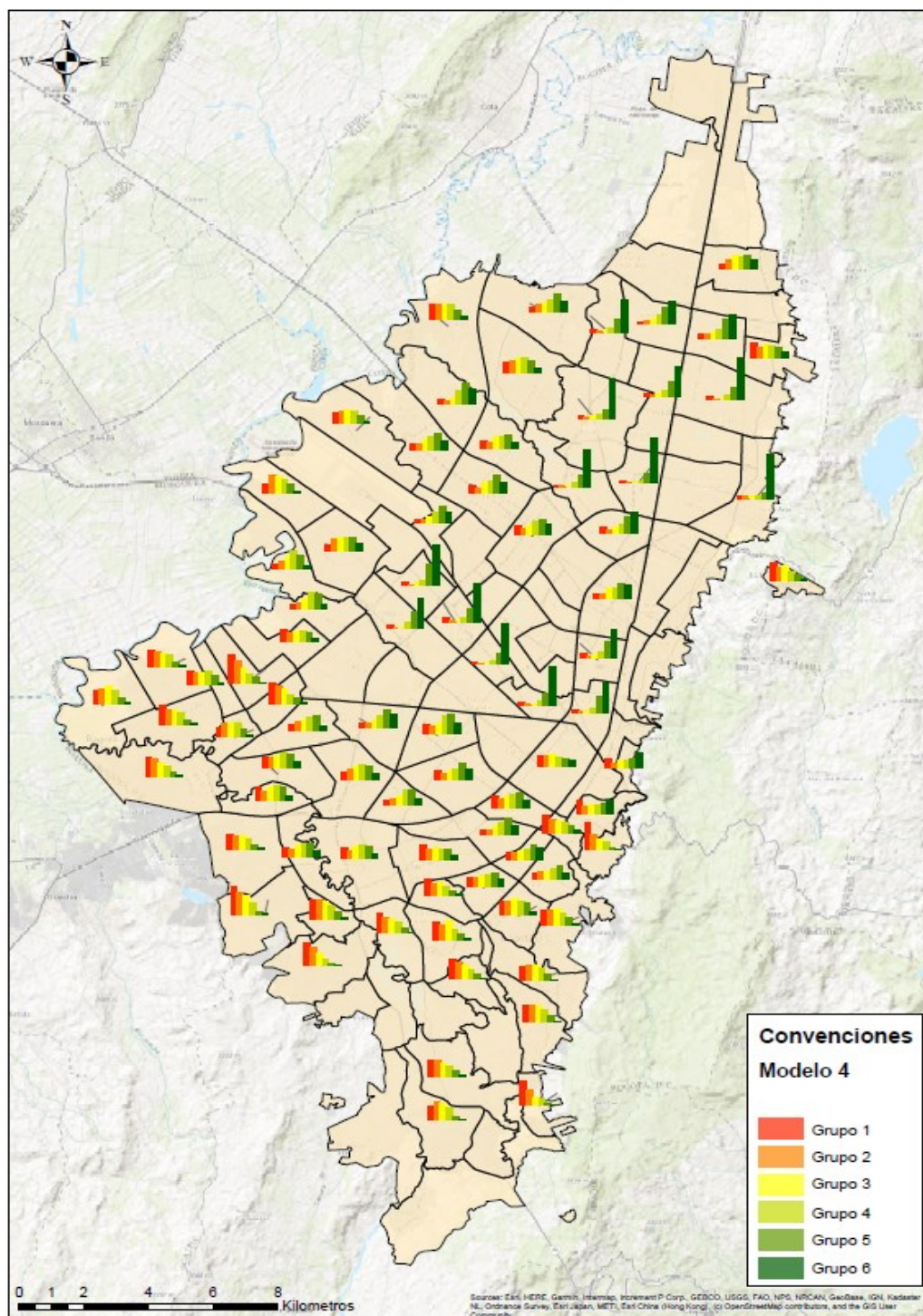
Ilustración 3. Mapa de distribución de grupos del índice de ingreso por UPZ. Modelo 3



Fuente: Elaboración propia.

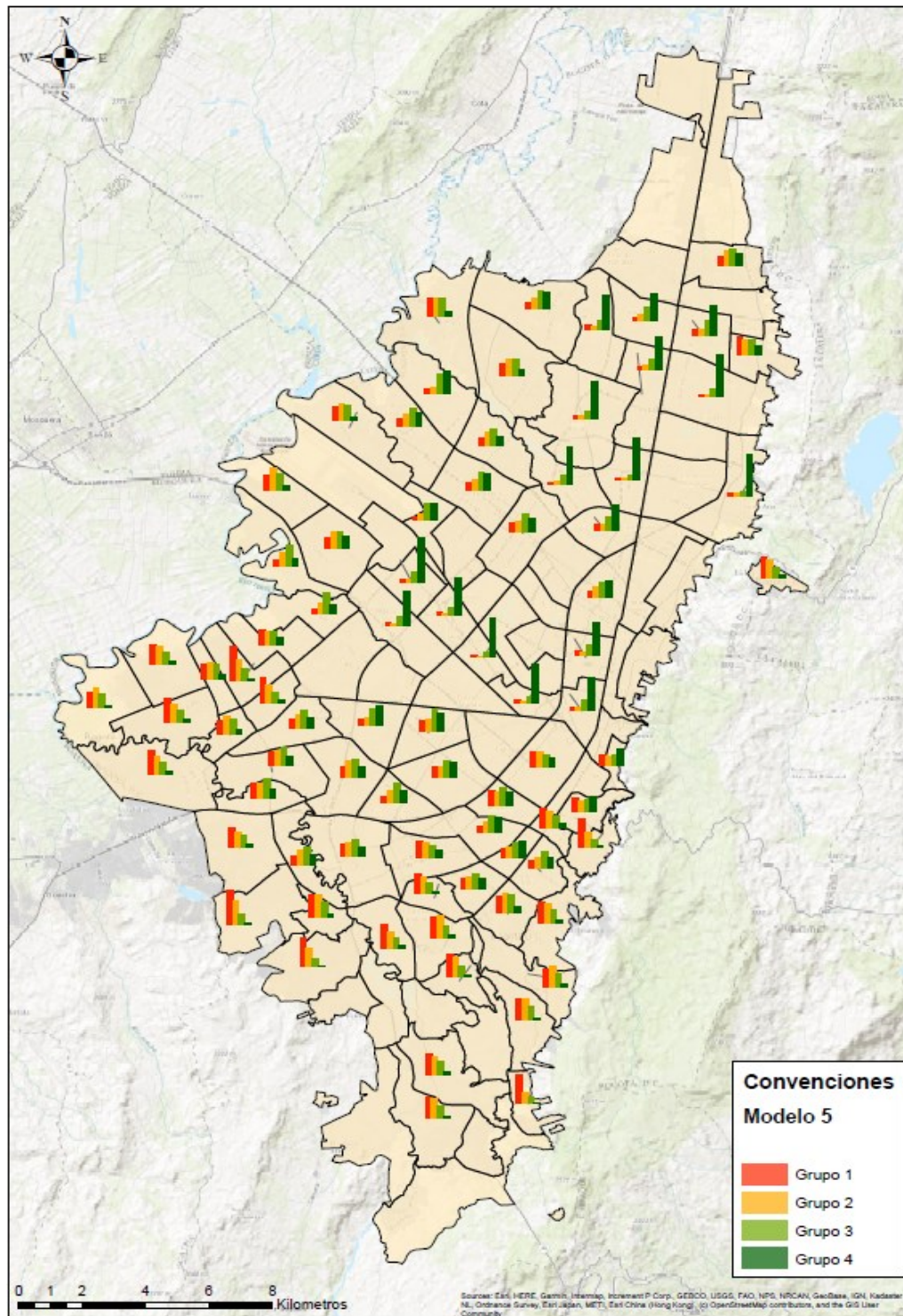


Ilustración 4. Mapa de distribución de grupos del índice de ingreso por UPZ. Modelo 4



Fuente: Elaboración propia.

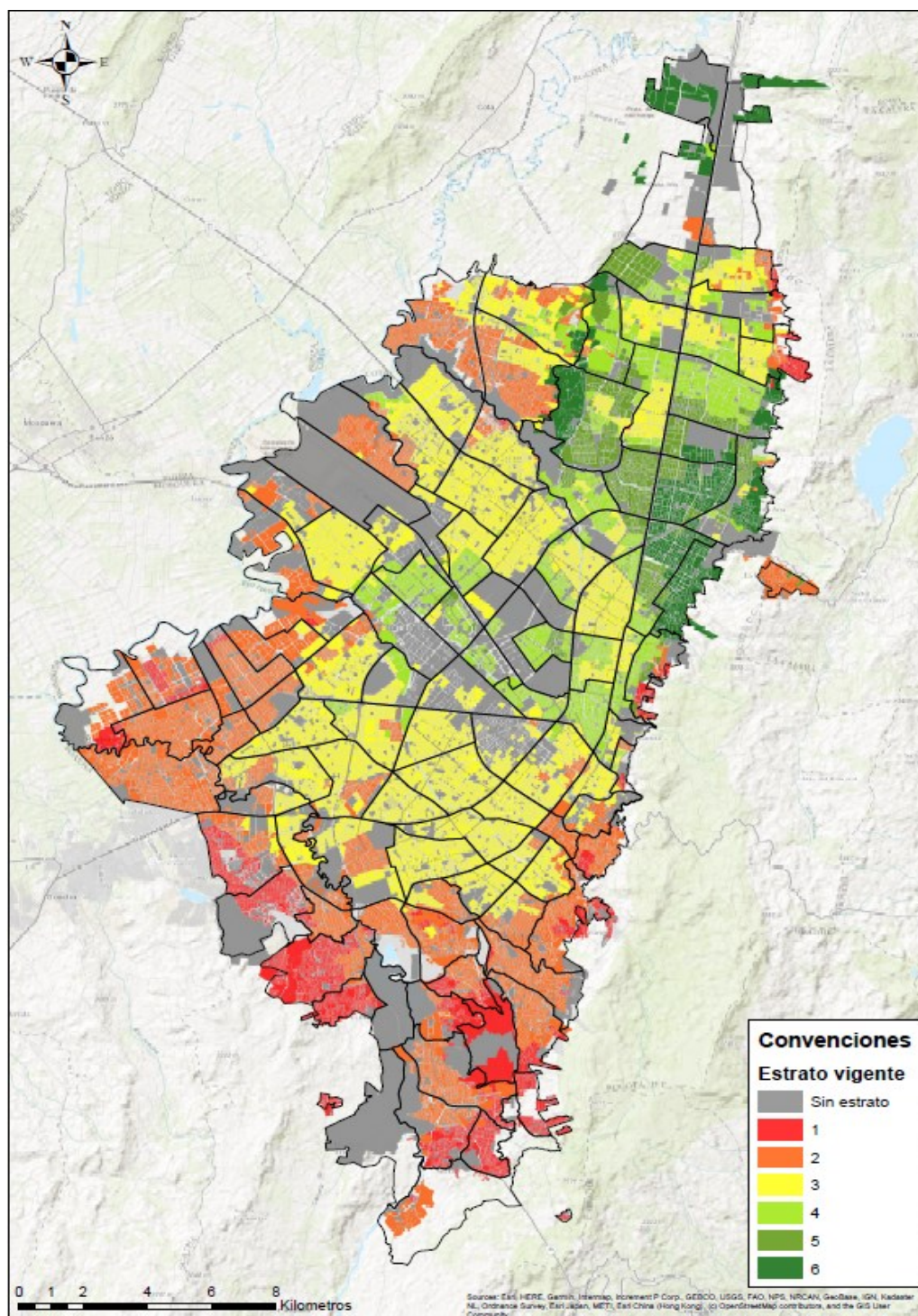
Ilustración 5. Mapa de distribución de grupos del índice de ingreso por UPZ. Modelo 5



Fuente: Elaboración propia.



Ilustración 6. Estratificación vigente por manzanas agrupadas por UPZ.



Fuente: Elaboración propia.

Se realiza un procedimiento adicional de estandarización de grupos para cada UPZ, seleccionando el grupo predominante de cada una y asignando su valor a todo el polígono, con la intención de obtener un resultado comparativo más sencillo de interpretar y de comparar con la estratificación vigente (Ilustraciones 7 a 11). Además, en la misma perspectiva, se incluye una ilustración con los rangos de valores promedios del índice por UPZ (ilustración 12). Es importante aclarar que desde la perspectiva de cartografía estos procedimientos descartan y desconocen gran cantidad de la información graficada, pues los polígonos de UPZ son demasiado grandes para obtener un resultado completo de la información procesada. Finalmente, con ayuda de la metodología IDW (Interpolación de Distancia Inversa Ponderada) se logra obtener un “mapa de calor” que permite inferir, con base en la información graficada, el grupo al que pertenecen aquellas UPZ que no contaban con información propia y que fueron listadas anteriormente, con el fin de obtener un mapa completo que arroje un panorama general de toda la ciudad.

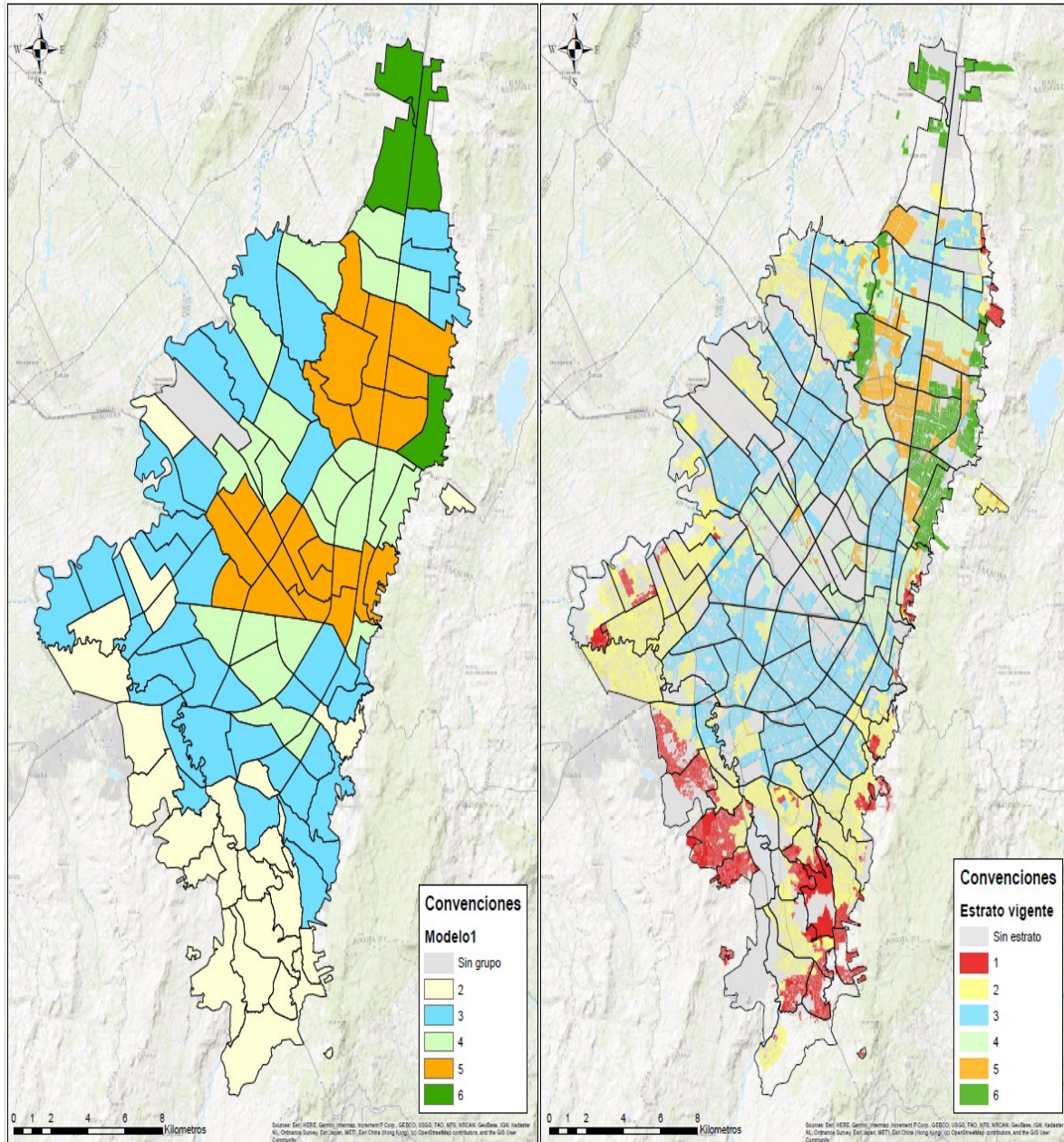
La metodología IDW supone explícitamente que los elementos cercanos entre sí son más parecidos que los que están más alejados. Para predecir un valor para cualquier ubicación sin medición, IDW usa los valores medidos que rodean a la ubicación de predicción. Los valores medidos más cercanos a la ubicación de predicción tienen más influencia en el valor predicho que los que están más alejados. IDW supone que cada punto medido tiene una influencia local que disminuye con la distancia. Otorga ponderaciones más altas a los puntos más cercanos a la ubicación de predicción y las ponderaciones disminuyen como una función de distancia (de ahí el nombre de distancia inversa ponderada). (tomado de: [Cómo funciona la interpolación de distancia inversa ponderada—ArcGIS Pro | Documentación](#)).

Con el criterio de proximidad que aplica la metodología IDW y bajo la presunción de que los grupos económicos tienden a agruparse espacialmente, se obtiene como resultado el modelamiento completo de los grupos económicos más representativos de toda la ciudad exceptuando la UPZ del aeropuerto El Dorado, que, a pesar de contar con un valor predicho en las interpolaciones, es ilógico asignarle dicho grupo al no contar con manzanas destinadas a la vivienda.

A continuación, se presenta una vista previa de los mapas comparativos entre grupos representativos de UPZ contra la estratificación vigente utilizando la misma paleta de colores para una fácil comparación:



Ilustración 7. UPZ por grupo predominante del modelo 1 obtenido por interpolación IDW vs. Estratificación vigente



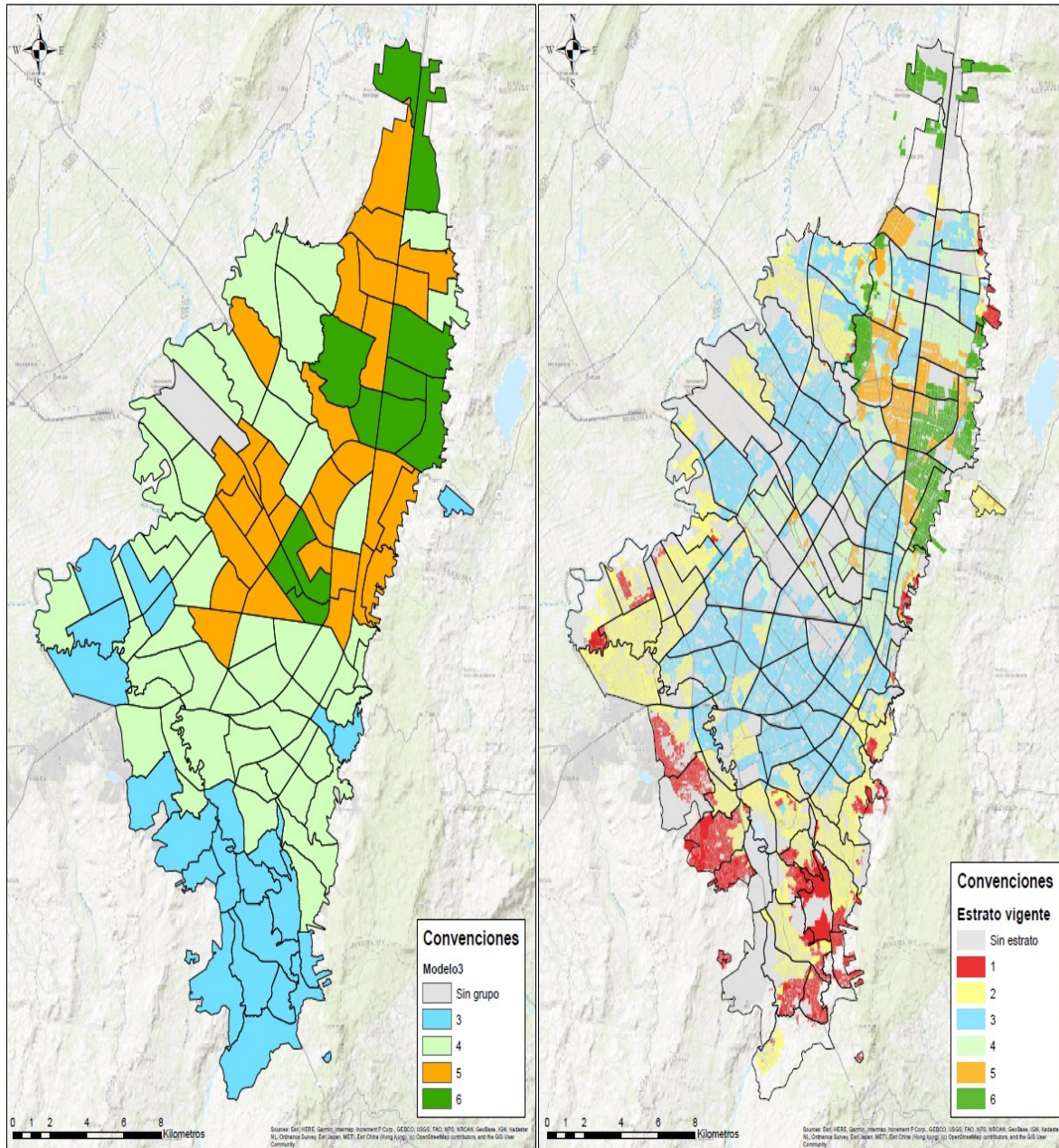
Fuente: Elaboración propia.







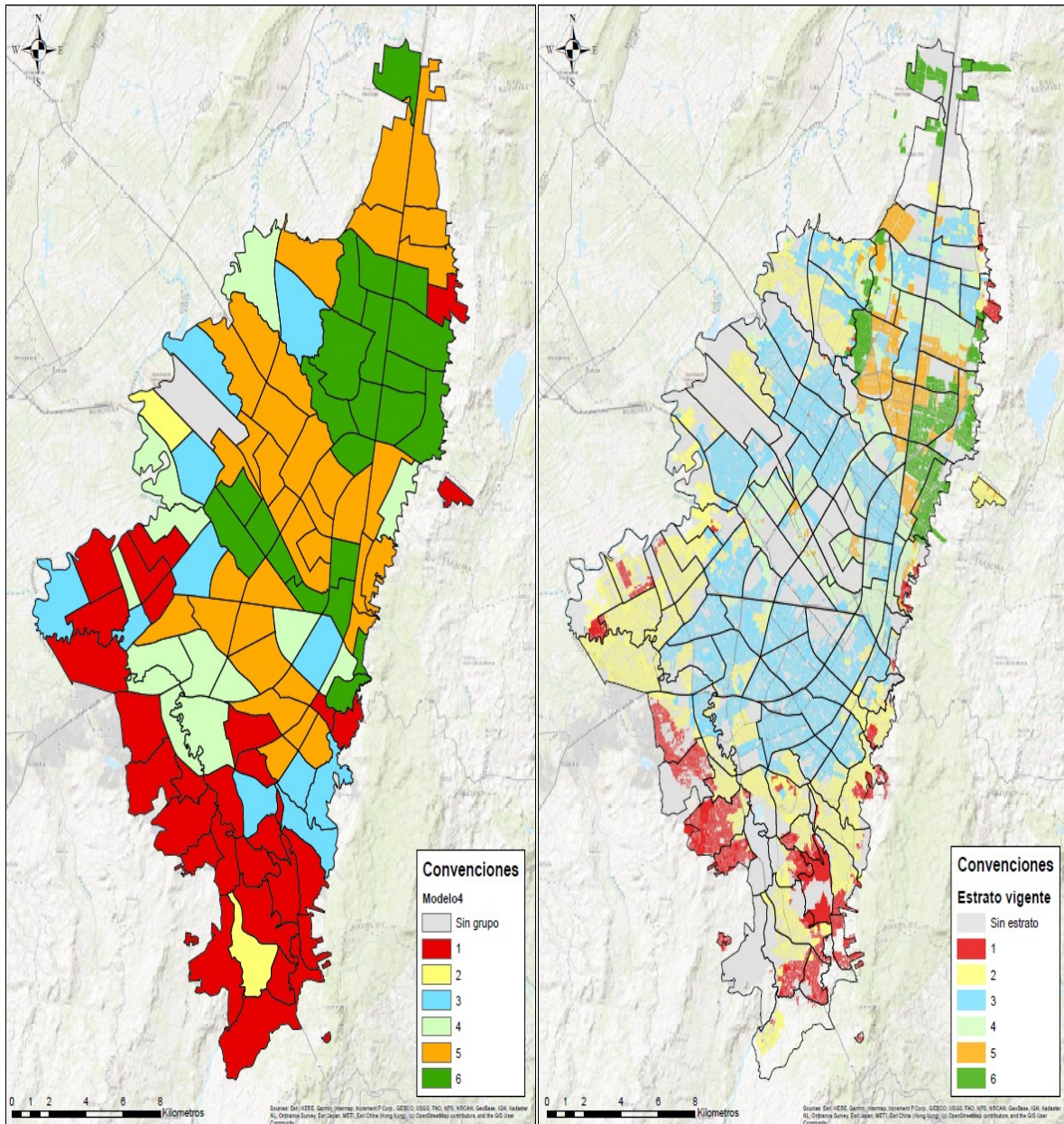
Ilustración 9. UPZ por grupo predominante del modelo 3 obtenido por interpolación IDW vs. Estratificación vigente



Fuente: Elaboración propia.



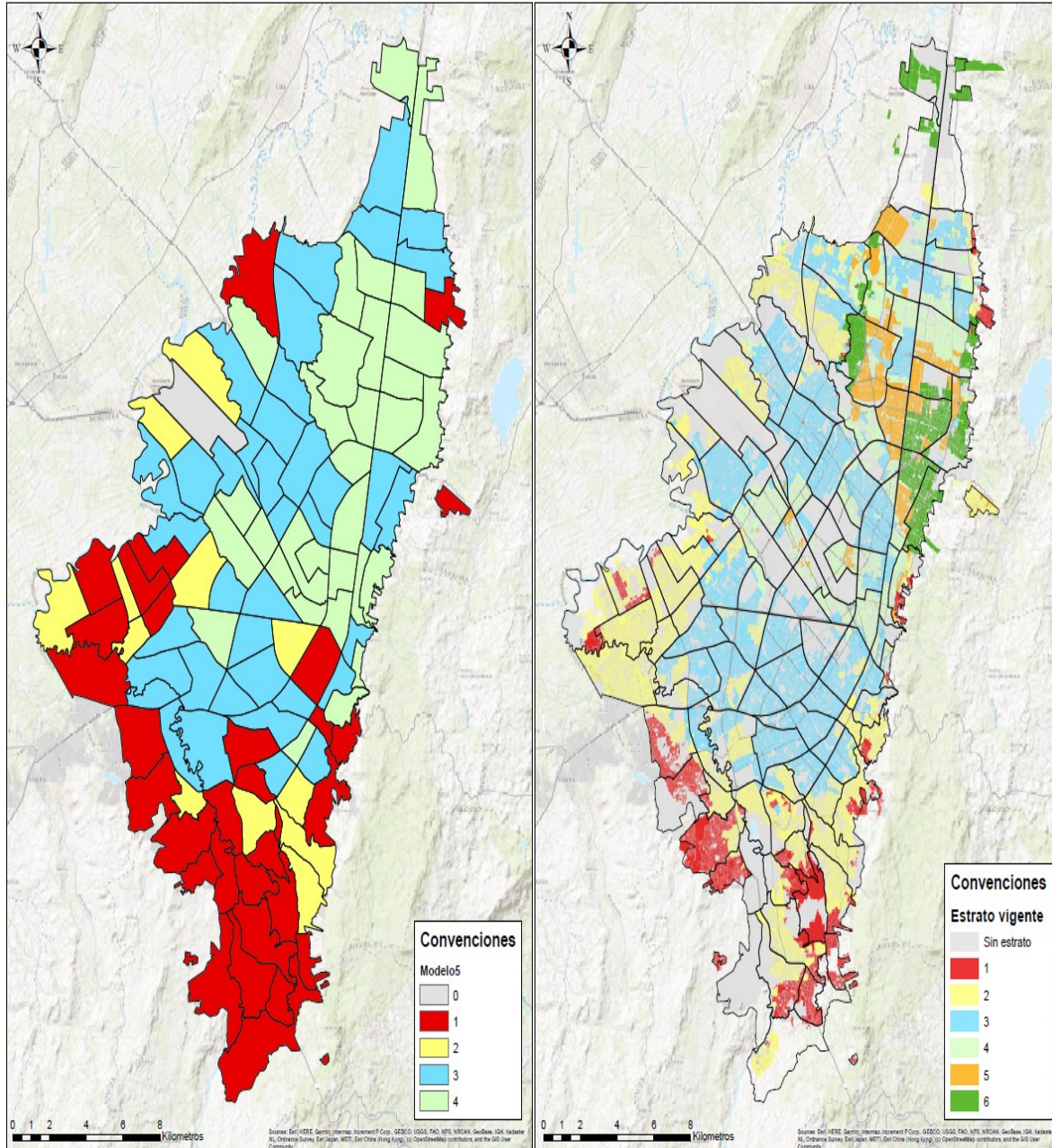
Ilustración 10. UPZ por grupo predominante del modelo 4 obtenido por interpolación IDW vs. Estratificación vigente



Fuente: Elaboración propia.

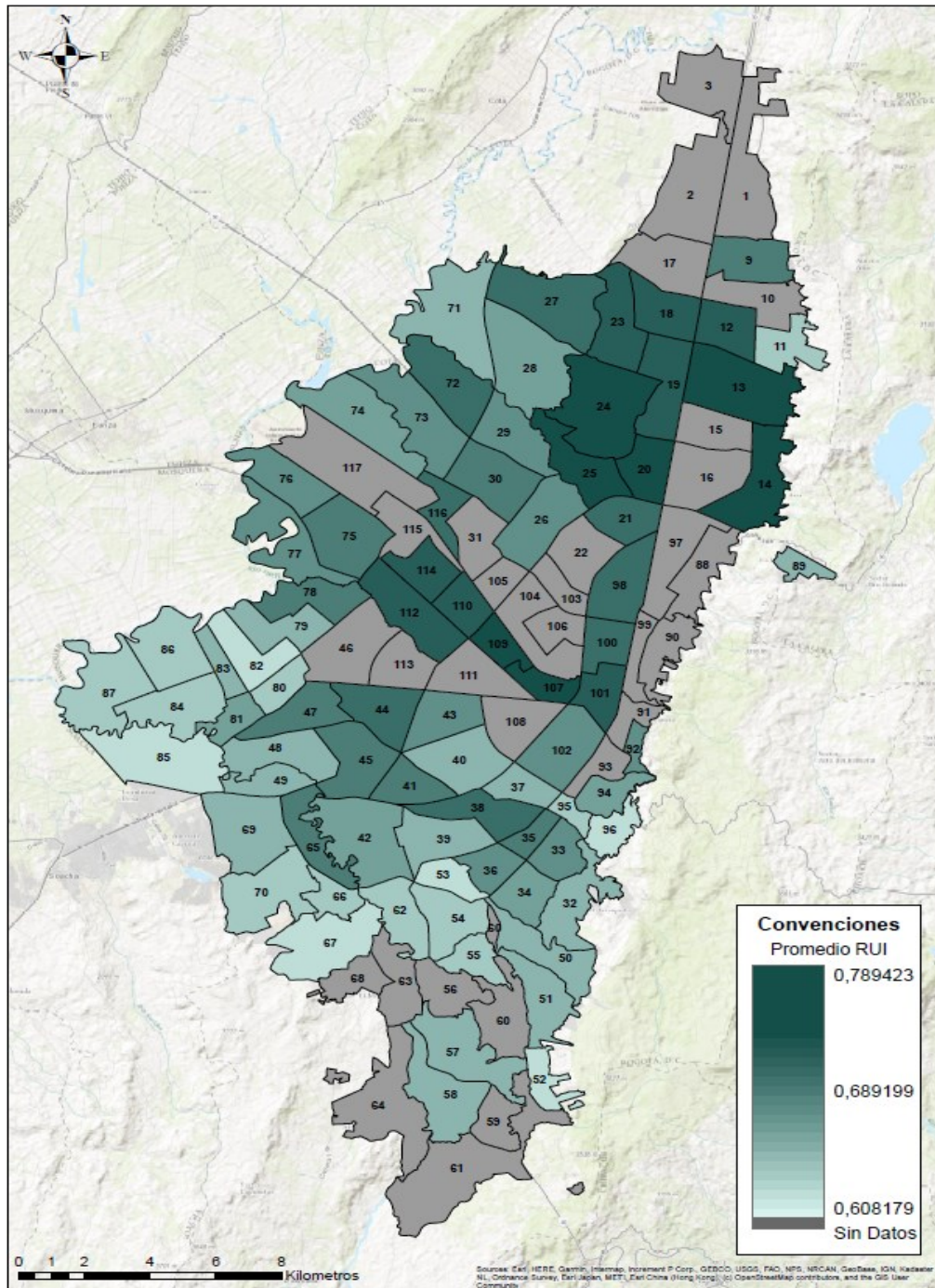


Ilustración 11. UPZ por grupo predominante del modelo 5 obtenido por interpolación IDW vs. Estratificación vigente



Fuente: Elaboración propia.

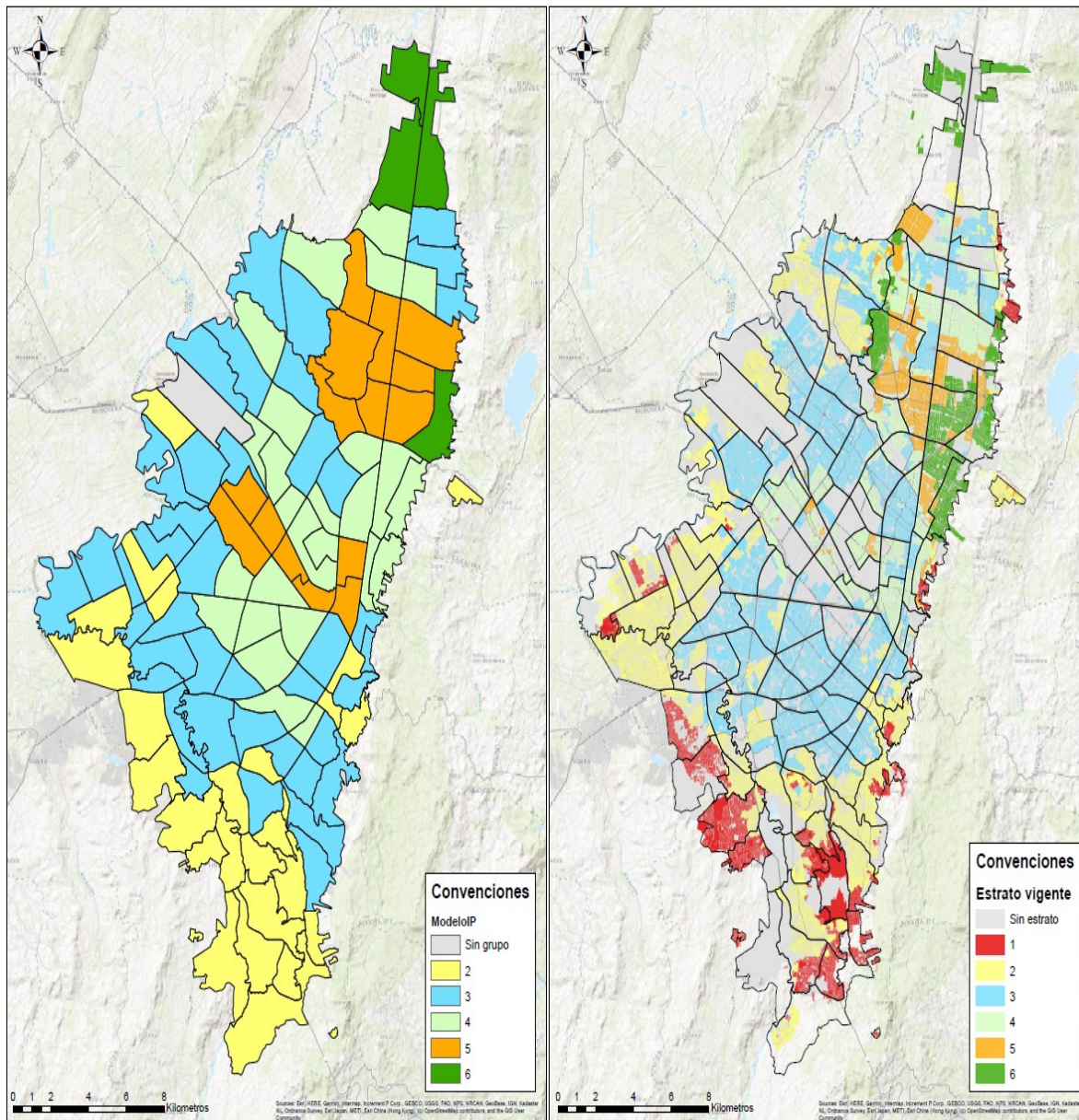
Ilustración 12. Valor promedio del índice de ingresos por UPZ



Fuente: Elaboración propia.



Ilustración 13. UPZ por grupo predominante del modelo con ingresos primarios obtenido por interpolación IDW vs. Estratificación



Fuente: Elaboración propia.

### 3.5 UPZ por estrato vs clasificación por índice RUI de ingresos Modelo 2

En las ilustraciones 7 a 11 y 13, el mapa de estratificación vigente presenta algunas UPZ parcialmente sin estrato en color gris como es el caso de las UPZ: 103, 104, 105, 108, 111, 112, 116. Sin embargo, según la información de la EMB 2021, en esas UPZ se encuentran hogares clasificados en algún estrato, tales los casos de las UPZ 112 y 104, en las cuales se encuentran hogares en estratos 3 y 4 y estratos 3, 4 y 5 respectivamente.

En un ejercicio comparativo con la clasificación por niveles de ingresos del Modelo 2 con aplicación del índice RUI, el resultado es que los hogares que según el mapa con la estratificación vigente aparecen en UPZ como sin estrato, quedan clasificados en los grupos: G2\_1, G2\_2, G2\_3 o G2\_4 (Cuadro 8).

*Cuadro 8. UPZ según estrato y clasificación por niveles de ingreso Modelo 2*

UPZ	Nombre UPZ	Estrato (tarifa de energía)						Modelo 2			
		0	1	2	3	4	5	G2_1	G2_2	G2_3	G2_4
112	Granjas de Techo				X	X		X	X	X	X
116	Álamos	X	X	X	X			X	X	X	X
104	TEUSAQUILLO: Parque Simón Bolívar - CAN + La Esmeralda				X	X	X	X	X	X	X
108-111	PUENTE ARANDA: Zona Industrial + Puente Aranda		X	X	X			X	X	X	X
103	BARRIOS UNIDOS: Parque Salitre + Doce de Octubre				X	X	X	X	X	X	X
105	ENGATIVÁ: Santa Cecilia + Jardín Botánico				X	X	X	X	X	X	X

*Fuente: Procesamiento propio de la EMB 2021 y elaboración propia*

Por otro lado, al comparar la estratificación vigente versus la clasificación del Modelo 2 se encuentran que algunos hogares pasarían de estrato 1 a clasificación G2\_3 como en el caso de la UPZ 116; en la UPZ 112, los hogares en estrato 3 y 4 quedarían clasificados según el Modelo 2 en los grupos G2\_1 o G2\_2 o G2\_3 o G2\_4 (Cuadro 9).

*Cuadro 9. UPZ según estrato y clasificación por niveles de ingreso Modelo 2*

*EMB 2021*

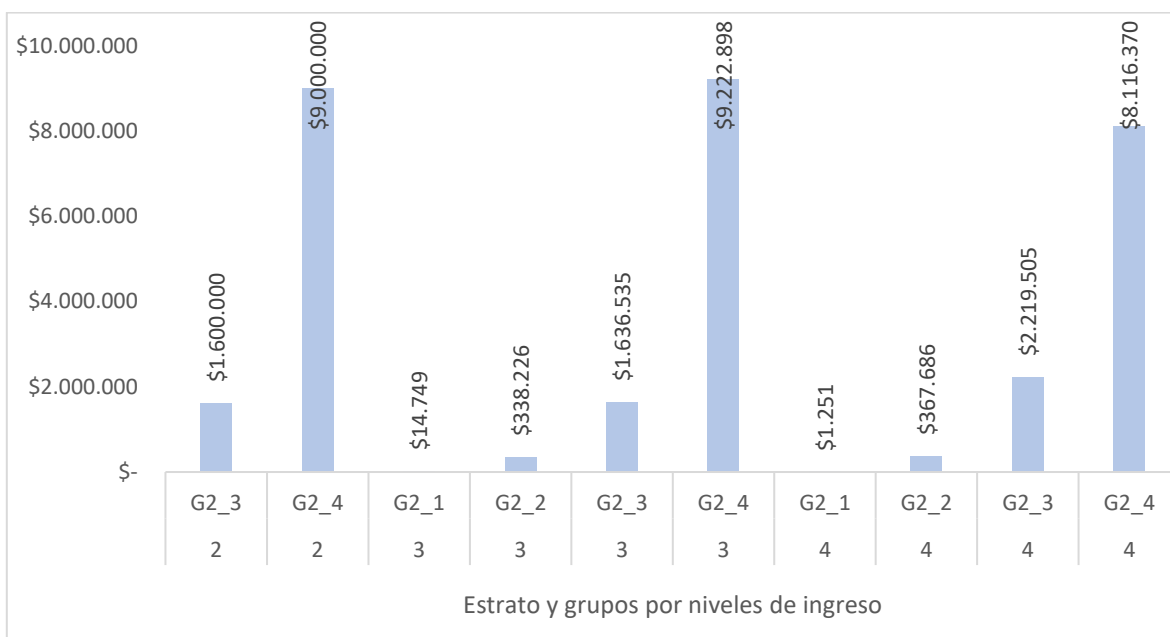
UPZ	Nombre UPZ	Estrato (tarifa de energía)	Modelo 2			
112	Granjas de Techo	3	X	X	X	X
		4	X	X	X	X
116	Álamos	0			X	
		1			X	
		2	X	X	X	
		3	X	X	X	X
104	TEUSAQUILLO: Parque Simón Bolívar - CAN + La Esmeralda	3	X		X	
		4	X	X	X	X
		5	X	X	X	X
108-111	PUENTE ARANDA: Zona Industrial + Puente Aranda	1	X	X	X	
		2	X	X	X	
		3	X	X	X	X
103	BARRIOS UNIDOS: Parque Salitre + Doce de Octubre	2			X	X
		3	X	X	X	X
		4	X	X	X	X
105	ENGATIVÁ: Santa Cecilia + Jardín Botánico	2	X			
		3	X	X	X	X
		4	X	X	X	X

*Fuente: Procesamiento propio de la EMB 2021 y elaboración propia*

Se presenta en el (Grafico 1 y 2) los ingresos promedios según clasificación del Modelo 2 contrastando con el estrato, con el objetivo de mostrar por ejemplo que en la que en la UPZ 103 en el estrato 2 se encuentra que el ingreso promedio en G2\_4 es de \$9.000.000 y en estrato 4 se encuentra que el ingreso promedio en G2\_2 es de \$ 367.686.

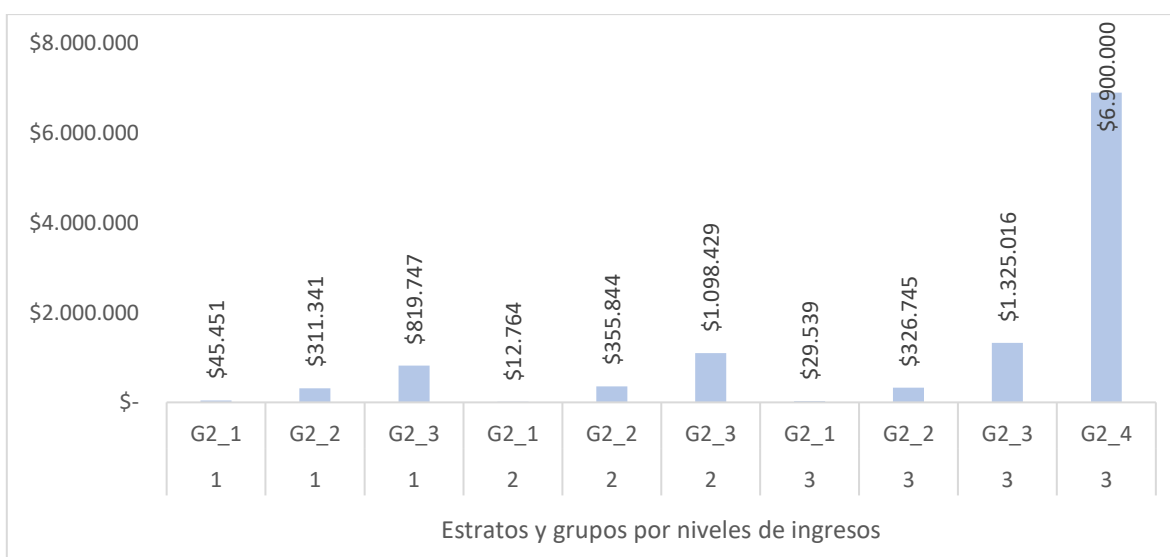
Comparando la UPZ 103 con las UPZ 108 y 111 se encuentran ingresos promedios mas bajos como es el caso en el estrato 2 se encuentra que el ingreso promedio en G2\_3 es de \$1.098.429 y el ingreso promedio más alto en estas UPZ 108 y 111 está en el estrato 3 en la clasificación G2\_4 es de \$6.900.000.

Gráfico 38. UPZ 103 Ingresos promedios por estrato (tarifa de energía) vs clasificación Modelo 2



Fuente: Para el estrato, EMB 2021, UPZ 103. Elaboración propia

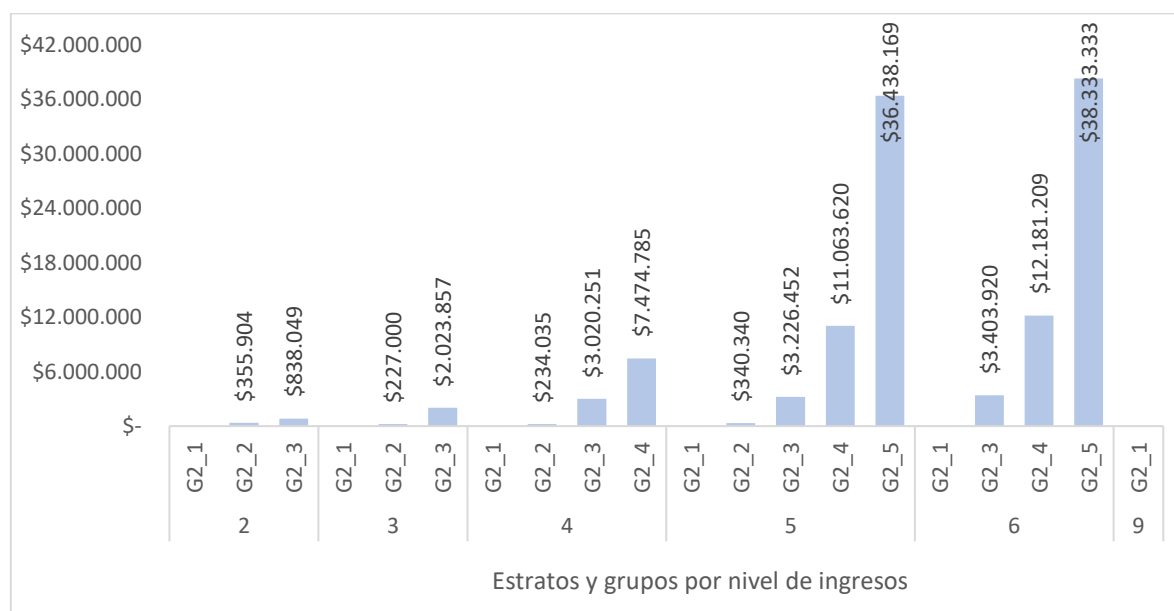
Gráfico 39. UPZ 108 y 111 Ingresos promedios por estrato (tarifa de energía) vs clasificación Modelo 2



Fuente: Para el estrato, EMB 2021, UPZ 108 y 111. Elaboración propia

Otro caso para mostrar: en estratos altos se encuentran hogares con ingresos promedios bajos, como en el caso de la UPZ 14: los hogares que actualmente están clasificados en estrato 5, estarían según la clasificación del índice RUI del Modelo 2 en el grupo G2\_2 con ingresos promedios bajos, por ejemplo de \$340.340; en el estrato 6 se encuentran hogares con ingresos promedios de \$3.403.920 clasificados en G2\_3 (Grafica 40).

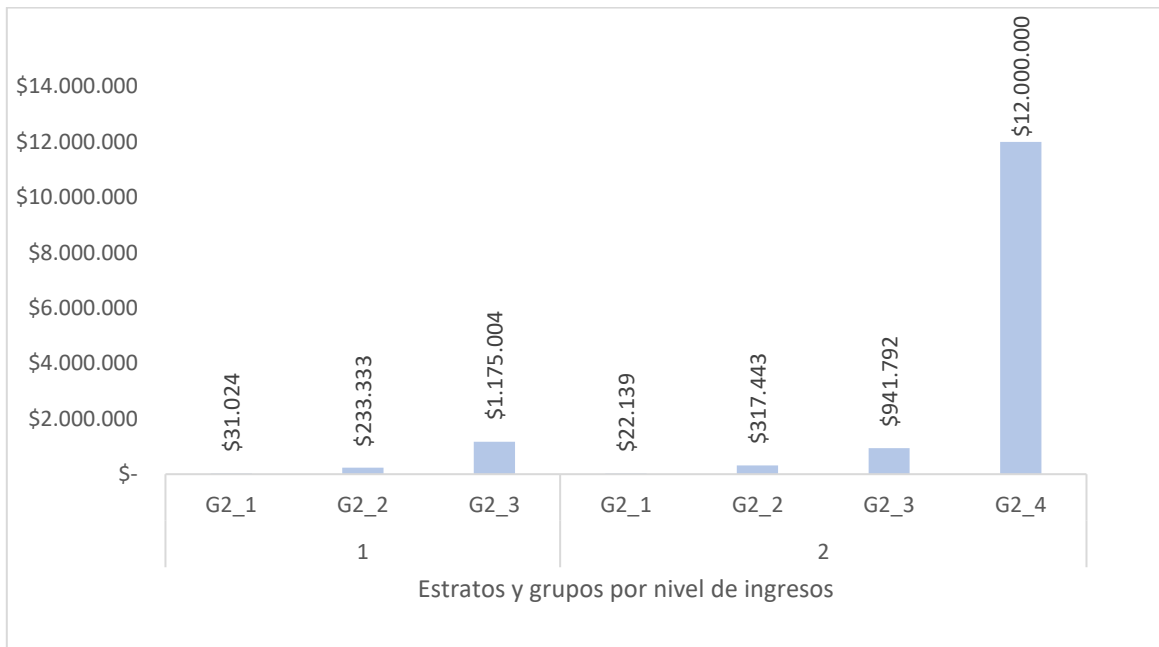
*Gráfico 40. UPZ 14 Ingresos promedios por estrato (tarifa de energía) vs clasificación Modelo 2*



*Fuente: Para el estrato, EMB 2021, UPZ 14. Elaboración propia*

En la UPZ 85 se encuentran hogares con ingresos promedios muy altos como es el caso de hogares clasificados en estrato 2 con la clasificación G2\_4 con ingresos promedios de \$12.000.000. (Grafico 41).

Gráfico 41. UPZ 85 Ingresos promedios por estrato (tarifa de energía) vs clasificación Modelo 2



Fuente: Para el estrato, EMB 2021, UPZ 85. Elaboración propia



#### 4. RELACIÓN CUANTITATIVA DEL ÍNDICE CONSTRUIDO CON LA ESTRATIFICACIÓN VIGENTE

En el cuadro 10 y los gráficos 42 y 43 se encuentra representada la distribución de la población de Bogotá por estratos socioeconómicos vigentes y grupos del modelo 2 del RUI, definidos con base en los ingresos secundarios.

Si se supone, solo en forma hipotética y en razón de tener un punto de comparación, que, para las tarifas de los servicios públicos domiciliarios, se asignan subsidios a los grupos de población pobre (1 y 2), no se fijan subsidios ni se imponen contribuciones a los grupos 3 y 4, y se establecen imposiciones a los grupos 5 y 6, se tendría un efecto de consideración en su clasificación y posiblemente en los montos de las tarifas de servicios públicos domiciliarios: el 68.6% de los hogares cambiarían de nivel en su clasificación, con modificación en su número de orden respecto a la estratificación socioeconómica vigente. Con el cambio del criterio de focalización el 50.6% de los hogares aumentarían de nivel (la población ubicada por encima de la diagonal del cuadro 10) y el 18% lo disminuiría (por debajo de la diagonal de ese cuadro).

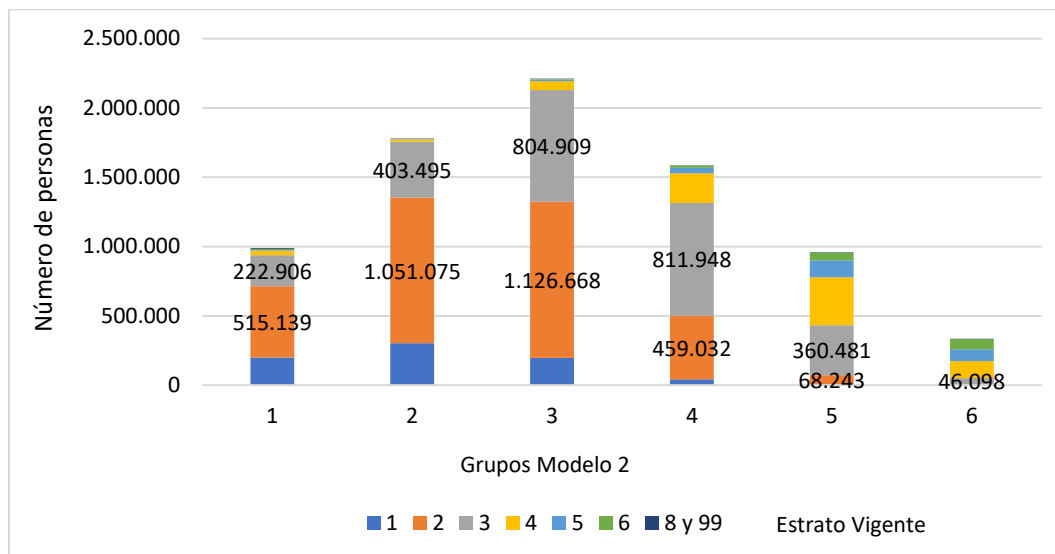
Al hacerlo, los grupos del modelo 2 quedarían conformados por una población heterogénea proveniente de distintos estratos. Como se observa en el gráfico 42, la población en hogares del estrato 1 se distribuirían entre los grupos 1, 2 y 3. Los del estrato 2, en forma mayoritaria, entre los grupos 1, 2 y 3, y los del 3 a 6, con proporciones variables, en los grupos 1 a 6. Este sería el proceso simulado que llevaría de la estratificación actual, con errores grandes de inclusión, a una situación donde quedarían suprimidos esos errores.

*Cuadro 10. Población por ubicación en los estratos socioeconómicos actuales y los grupos del modelo 2 del ingreso secundario del RUI- 2021.*

Estrato Energía	Grupos Modelo 2						Total
	1	2	3	4	5	6	
1	198.736	302.524	197.506	42.907	3.899	322	745.894
2	515.139	1.051.075	1.126.668	459.032	68.243	5.659	3.225.816
3	222.906	403.495	804.909	811.948	360.481	46.098	2.649.837
4	32.072	15.251	63.396	215.730	345.902	122.399	794.750
5	6.588	3.153	11.888	37.424	121.458	84.380	264.891
6	5.275	2.081	5.154	16.120	59.110	74.187	161.927
8 y 99	8.449	3.569	1.817	301	132	71	14.339
Total	989.165	1.781.148	2.211.338	1.583.462	959.225	333.116	7.857.454

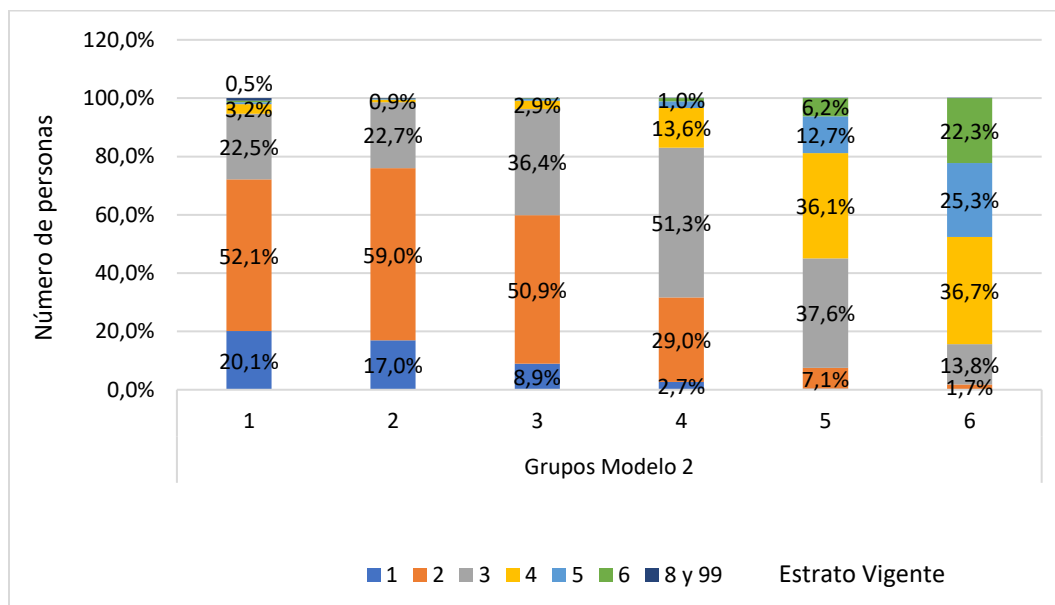
*Fuente: procesamiento propio de la EMB 2021.*

Gráfico 42. Distribución de la población de Bogotá por grupos del Modelo 2 y la estratificación vigente- 2021



Fuente: procesamiento propio de la EMB 2021.

Gráfico 43. Distribución porcentual de la población de Bogotá por grupos del Modelo 2 y la estratificación vigente- 2021



Fuente: procesamiento propio de la EMB 2021.

## 5. ESTIMACIÓN DE LOS ERRORES DE INCLUSIÓN Y DE EXCLUSIÓN

Se presentan en esta sección, retomado lo expuesto en el documento del producto tres, los errores de inclusión y exclusión para la estratificación vigente y para cada uno de los modelos elaborados, tomando como referencia tanto la pobreza extrema monetaria como la no extrema. Se presenta un resumen de lo expuesto sobre este tema en el documento 3 del proyecto.

El Cuadro 11 contiene el cálculo de los errores de inclusión y exclusión de la estratificación actual y de los agrupamientos de hogares resultantes de los cinco modelos planteados en el documento del producto 3.

*Cuadro 11. Errores de inclusión y de exclusión de la estratificación vigente y los grupos de los modelos 1 a 5 – Bogotá – 2021*

<b>Pobreza monetaria no extrema</b>		
	<b>Error de inclusión</b>	<b>Error de exclusión</b>
<i>Estratificación actual</i>	77,24	2,34
<i>Modelo 1</i>	0	0
<i>Modelo 2</i>	0	0
<i>Modelo 3</i>	0.200	2.44
<i>Modelo 4</i>	22,79	0
<i>Modelo 5</i>	22,79	0
<b>Pobreza monetaria extrema</b>		
	<b>Error de inclusión</b>	<b>Error de exclusión</b>
<i>Estratificación actual</i>	82.84	4.48
<i>Modelo 1</i>	25.93	0
<i>Modelo 2</i>	25.93	0
<i>Modelo 3</i>	25.93	0
<i>Modelo 4</i>	42,82	0
<i>Modelo 5</i>	42,82	0

*Fuente: Cálculos propios con base en la EMB 2021.*

Se acepta que la estratificación vigente y las agrupaciones de ingresos deben permitir diferenciar a los segmentos que merecen recibir subsidios en razón de su condición de pobreza o pobreza extrema. En el caso de la estratificación estos segmentos se ubicarían en los estratos 1, 2 y 3. En el modelo 1, en los grupos 0, 1 y 2; en los modelos 2, 3 y 5, en los grupos 1 y 2; y en el modelo 4, en los grupos 1, 2 y 3.

El ejercicio de estimación de dichos errores se hace con base en la información de la EMB 2021, que se toma como laboratorio, asumiendo el supuesto de que no hay errores en la información de ingresos registrada en esa encuesta.

Los errores de inclusión expresan, el porcentaje de población con ingresos superiores a las líneas de pobreza e indigencia, que pertenecen a los estratos o a los grupos a los que se asignaría un subsidio. Muestran la magnitud del error de clasificar como pobres, y con derecho a subsidios, a quienes no lo son.

Los errores de exclusión expresan la proporción de población pobre o pobre extrema a la que no se clasifica como merecedora de un subsidio en razón de su ubicación en la estratificación vigente o en los grupos de ingreso de los modelos.

El resultado de los ejercicios realizados con la información de la EMB 2021 lleva a concluir que los modelos con mejores resultados son el 1 y 2, que serían los recomendados. El modelo 2 tiene una ventaja adicional frente al 1, al diferenciar el segmento de unidades de gasto sin ingresos.

Es de resaltar la apreciable mejora que representaría la implantación de una focalización basada en escalas de ingresos, que corregiría el alto nivel del error de inclusión de la actual estratificación, en su aplicación a las tarifas de los servicios públicos domiciliarios, que se sitúa en el 77%.

La ventaja de los modelos 1 y 2, frente a los otros modelos, se encuentra precisamente en que, con información insesgada, suprimirían los errores de inclusión, además de los de exclusión, que en los modelos 4 y 5, respecto a la pobreza no crítica, se ubicaría en torno al 22%. Para la pobreza crítica reduciría en casi 40% estos errores frente a los modelos 4 y 5.

## 6. LOS REGISTROS ADMINISTRATIVOS (RA) DISPONIBLES Y EL REGISTRO UNIVERSAL DE INGRESOS (RUI).

En la sección se exponen los ejercicios realizados para integrar la información de los dos RA de que se dispuso en el proyecto, en la construcción del RUI. Sobre la información de la PILA se adelantaron dos ejercicios. Uno tendiente a cotejar la información proveniente de esa fuente sobre número de trabajadores y valor de las remuneraciones laborales, con la obtenida de la EMB 2021. El otro dirigido a caracterizar algunos grupos de trabajadores identificables en la PILA, con la información de esa misma encuesta para poblaciones laborales equivalentes. Respecto al SISBEN se llevó a cabo un resumido diagnóstico sobre la calidad de su información sobre los ingresos personales y de los hogares; y, al comprobar que su calidad es deficiente para los propósitos del RUI, se acometió el ejercicio de hacer una estimación de los ingresos de los hogares encuestados en el SISBEN, a partir de características socioeconómicas, con base en los parámetros obtenidos en un modelo de regresión aplicado en la misma EMB 2021. Se intenta de esta forma obtener orientaciones prácticas sobre la forma de utilización de los RA en la elaboración del índice de ingresos. En particular, es explorada por esta vía práctica una de las tres fuentes de información del RUI que es mencionada en el proyecto de decreto del DNP sobre este Registro<sup>17</sup>: la que tiene origen en registros administrativos en dimensiones diferentes al ingreso que permita establecer las características socioeconómicas del hogar, en ausencia de la información del ingreso observado.

### 6.1. Las diferencias entre la información de la PILA y la de la EMB 2021 sobre empleo y aportes a la seguridad social en salud

Con la información de la PILA el ejercicio consistió en comparar la información provista por este RA, relativa al número de cotizantes y a los promedios de remuneración corriente laboral, con la proveniente de la EMB 2021. Para llevarlo a cabo se seleccionó, dentro de las bases PILA, lo correspondiente a 2021 para que coincidieran los periodos de referencia de las dos fuentes. Además, se obtuvo la agregación en esas bases de los registros de cada persona, de acuerdo con su número de identificación registrado<sup>18</sup> en la base de datos, para obtener su remuneración mensual, media laboral a lo largo del año, y se seleccionaron en la encuesta únicamente los afiliados a los regímenes contributivo, especiales y de excepción.

La importancia de avanzar en la evaluación de la información de la PILA reside en la alta proporción de los ingresos de las personas que provienen de su trabajo y la elevada cobertura que tiene el aseguramiento en salud. El aporte de la información provista por los

---

<sup>17</sup> Véase al respecto el numeral 4.3.5 y el anexo 2 del documento del producto 3 del proyecto en el que se ubica el presente texto.

<sup>18</sup> Es de advertir, sin embargo, que no se contó con los números de identificación completos de los cotizantes registrados en la PILA.

distintos sistemas de seguridad social constituye uno de los elementos fundamentales para construir el RUI.

El Ingreso Base de Liquidación (IBL) declarado por los asalariados en la PILA se equiparó al ingreso corriente laboral registrado en la Encuesta, sin tomar en cuenta los ingresos correspondientes a primas, vacaciones y otros conceptos que generalmente no son declarados para la liquidación de aportes al régimen contributivo de salud y los regímenes especiales y de excepción. Del IBL de los trabajadores independientes se obtuvo una estimación del ingreso bruto (multiplicando el IBL por 2.5<sup>19</sup>) y luego descontándole a ese resultado el 40%, que correspondería a los impuestos y demás gastos en que se incurre con el desarrollo de la actividad laboral. Se fijó, así, una aproximación al ingreso neto de estos trabajadores que es el considerado en la encuesta para ellos.

En el Cuadro 12 se expone el cotejo realizado entre las estimaciones de las dos fuentes. En cuanto al número de trabajadores, que se obtiene de los cotizantes en la PILA, se destaca que esta fuente presenta aparentemente una sobreestimación, tomando como referencia la información de la encuesta. La PILA registra un número de cotizantes afiliados a este régimen, superior en 1.600.000 al estimado de trabajadores con la encuesta y este valor supera el total de trabajadores que calcula la encuesta para el conjunto de Bogotá. En todos los otros ámbitos de comparación la PILA rebasa lo obtenido de la encuesta. Solo en el número de cotizantes trabajadores independientes la diferencia entre los segmentos comparados es inferior al 10%. En cuanto al número de asalariados cotizantes en el régimen subsidiados la diferencia entre la PILA y la EMB es de 72%.

Esta situación lleva a concluir que hay “duplicación” en el conteo de un número apreciable de cotizantes en la PILA, o el periodo de cotización no coincide con el de trabajo, ya que no es probable que en la encuesta haya un subregistro de trabajadores en la magnitud de lo que lo estaría mostrando su comparación con PILA. Para controlar esos registros duplicados sería necesario contar con las claves de identificación completas de los trabajadores, con las cuales no se dispuso. Y hacer igualmente una revisión de la compatibilidad de las categorías de clasificación de cada una de las dos fuentes.

En cuanto a la información sobre el IBL se evidencian discrepancias entre las dos fuentes en cuanto al promedio de las remuneraciones. En algunos grupos las diferencias son grandes. En los afiliados a los regímenes especiales son de más del 200%. En los empleados domésticos de 21.4% y en las categorías de asalariados e independientes se sitúan en niveles más moderados de -4.2 y 1.8% respectivamente.

En la perspectiva de ir consolidando los pilares básicos del RUI convendría ir regularizando la producción de la información provista por ese RA y ajustando las categorías y forma de procesamiento para lograr un registro adecuado y compatible de esa información, con la participación del DANE y el DNP.

Este ejercicio ilustra la compleja tarea que significa la adecuación de los RA, para su integración coherente en el RUI. Y señala cómo los resultados obtenidos impiden la utilización de la PILA, en su conformación actual disponible, para obtener un estimado

---

<sup>19</sup> En la PILA los independientes declaran el 40% de sus ingresos laborales brutos como IBL.

confiable de los ingresos laborales de los cotizantes al Sistema de Seguridad Social en Salud.

*Cuadro 12. Información sobre número de cotizantes y trabajadores y remuneración laboral de la PILA y la EMB 2021. Bogotá-2021*

Concepto de comparación	PILA	EMB 2021	Diferencia	Diferencia (%)
<b>Número de cotizantes (PILA) y de trabajadores cotizantes (EPM)</b>				
Total cotizantes y trabajadores	-	3.541.646	-	
Total trabajadores cotizantes al régimen contributivo	4.358.349	2.749.610	1.608.739	58,5%
Total trabajadores asalariados, independientes y domésticos afiliados al régimen contributivo y a los regímenes especiales y de excepción	4.168.182	2.768.843	1.399.339	50,5%
Número asalariados régimen contributivo	3.354.244	1.951.175	1.403.069	71,9%
Número independientes régimen contributivo	758.351	718.238	40.113	5,6%
Número empleados domésticos afiliados al régimen contributivo	45.465	31.040	14.425	46,5%
Número trabajadores afiliados a los regímenes especiales y de excepción	10.122	68.390	-58.268	-85,2%
<b>Remuneración laboral promedio mensual (pesos corrientes)</b>				
Remuneración laboral promedio régimen contributivo	-	2.096.808	-	
Remuneración laboral promedio régimen especial	8.013.365	2.637.526	5.375.840	203,82%
Promedio remuneración asalariados	1.978.855	2.064.764	-85.909	-4,16%
Promedio remuneración de independientes (1)	1.849.846	1.816.493	33.353	1,84%
Promedio remuneración domésticos	932.415	768.010	164.405	21,41%

(1) 40% del ingreso bruto laboral.

*Fuente: elaboración propia a partir de un procesamiento especial de PILA y la EMB 2021.*



## 6.2. Caracterización externa de los grupos de trabajadores identificados en la PILA

Un acercamiento a la caracterización de los segmentos de trabajadores que cubre la PILA puede obtenerse a través de fuentes estadísticas externas que tengan las variables necesarias para identificar esos grupos de trabajadores, y que ofrezcan información de interés para conocer sus características. La prueba realizada en esta perspectiva recurrió también a la EMB 2021. Se llevó a cabo para los mismos cuatro tipos de trabajadores considerados en la comparación relativa a número de cotizantes y trabajadores y a remuneraciones (cuadro 12): asalariados, independientes y trabajadores domésticos afiliados al régimen contributivo, y trabajadores afiliados a los regímenes especiales y de excepción. Y se toman como ejemplos de indicadores de caracterización los cinco que anteriormente se seleccionaron para ilustrar las condiciones de vulnerabilidad de los hogares (numeral 3.2).

Se presentan en los cuadros siguientes (13 a 16) los resultados del ejercicio para cada uno de los cuatro grupos de trabajadores. La información contenida en ellos podría ser utilizada como complemento para definir un criterio de focalización *mixto*<sup>20</sup> que tomara en cuenta información socioeconómica. Ayuda, además, a calificar comparativamente la situación de los distintos agrupamientos. En la información presentada son las trabajadoras domésticas la que se encuentran en mayor desventaja. Y son los afiliados a regímenes especiales o de excepción los que cuentan con mayor ventaja.

---

<sup>20</sup> Véase el numeral 5.3.1. (Ampliación del índice con información socioeconómica sobre vulnerabilidad y privación) del documento del producto 3 del proyecto.

*Cuadro 13. Indicadores de vulnerabilidad de los asalariados afiliados al régimen contributivo de salud. Porcentajes de personas pertenecientes a hogares con la característica del indicador dentro de cada grupo de ingresos. Bogotá-2021*

Grupos modelo 2	Nivel de ingreso	Porcentaje de personas por rango de ingreso				
		Hogares con menores de edad	Hogares con mayores de 60 años sin ingreso	Hogares con personas con discapacidad	Hogares sin tenencia propia	Hogares en hacinamiento
G2_1	Menor a 196.583	40,88%	22,36%	21,90%	50,46%	0,69%
G2_2	196.583,01 a 476.133	51,59%	21,59%	39,35%	55,41%	2,31%
G2_3	476.133,01 a 6.638.333	49,97%	18,41%	26,00%	56,58%	1,36%
G2_4	6.638.333,01 a 29.800.000	45,89%	12,26%	13,51%	37,18%	0,36%
G2_5	29.800.000,01 a 70.000.000	51,71%	13,80%	11,55%	40,05%	0,00%
G2_6	Más de 70.000.000					

*Fuente: elaboración propia a partir de un procesamiento especial de PILA y la EMB 2021.*

*Cuadro 14. Indicadores de vulnerabilidad de los trabajadores independientes afiliados al régimen contributivo de salud. Porcentajes de personas pertenecientes a hogares con la característica del indicador dentro de cada grupo de ingresos. Bogotá-2021*

Grupos modelo 2	Nivel de ingreso	Porcentaje de personas por rango de ingreso				
		Hogares con menores de edad	Hogares con mayores de 60 años sin ingreso	Hogares con personas con discapacidad	Hogares sin tenencia propia	Hogares en hacinamiento
G2_1	Menor a 196.583	35,97%	34,03%	26,72%	45,36%	0,64%
G2_2	196.583,01 a 476.133	46,23%	34,74%	36,76%	51,32%	2,46%
G2_3	476.133,01 a 6.638.333	43,48%	25,24%	23,84%	51,07%	0,66%
G2_4	6.638.333,01 a 29.800.000	37,25%	19,31%	16,82%	39,82%	0,10%
G2_5	29.800.000,01 a 70.000.000	48,12%	27,84%	10,69%	36,37%	0,00%
G2_6	Más de 70.000.000	28,44%	46,87%	27,12%	32,86%	0,00%

*Fuente: elaboración propia a partir de un procesamiento especial de PILA y la EMB 2021.*

*Cuadro 15. Indicadores de vulnerabilidad de los trabajadores domésticos afiliados al régimen contributivo de salud. pertenecientes a hogares con la característica del indicador dentro de cada grupo de ingresos. Bogotá-2021*

Grupos modelo 2	Nivel de ingreso	Porcentaje de personas por rango de ingreso				
		Hogares con menores de edad	Hogares con mayores de 60 años sin ingreso	Hogares con personas con discapacidad	Hogares sin tenencia propia	Hogares en hacinamiento
G2_1	Menor a 196.583	45,04%	47,58%	39,37%	36,10%	3,05%
G2_2	196.583,01 a 476.133	50,57%	31,50%	38,08%	59,00%	3,50%
G2_3	476.133,01 a 6.638.333	49,67%	24,26%	28,40%	64,96%	3,03%
G2_4	6.638.333,01 a 29.800.000					
G2_5	29.800.000,01 a 70.000.000					
G2_6	Más de 70.000.000					

*Fuente: elaboración propia a partir de un procesamiento especial de PILA y la EMB 2021.*

*Cuadro 16. Indicadores de vulnerabilidad de los afiliados a los regímenes especiales y de excepción de salud. pertenecientes a hogares con la característica del indicador dentro de cada grupo de ingresos. Bogotá-2021*

Grupos modelo 2	Nivel de ingreso	Porcentaje de personas por rango de ingreso				
		Hogares con menores de edad	Hogares con mayores de 60 años sin ingreso	Hogares con personas con discapacidad	Hogares sin tenencia propia	Hogares en hacinamiento
G2_1	Menor a 196.583	46,81%	20,70%	19,39%	48,36%	0,00%
G2_2	196.583,01 a 476.133	48,49%	35,08%	48,09%	49,03%	3,56%
G2_3	476.133,01 a 6.638.333	50,02%	17,75%	26,73%	44,32%	0,52%
G2_4	6.638.333,01 a 29.800.000	42,90%	16,96%	6,92%	24,79%	2,58%
G2_5	29.800.000,01 a 70.000.000	100,00%	0,00%	0,00%	32,75%	0,00%
G2_6	Más de 70.000.000	0,00%	100,00%	100,00%	0,00%	0,00%

*Fuente: elaboración propia a partir de un procesamiento especial de PILA y la EMB 2021.*

## 6.3. La información sobre ingresos del SISBÉN IV

### 6.3.1. La información sobre ingresos

El otro RA de que se dispuso, que podría aportar información sobre los ingresos de los individuos y los hogares, es el SISBÉN IV. Se hizo una revisión de lo provisto por esta fuente para poner a prueba la confiabilidad de esa información, en la perspectiva de la construcción del RUI. Además, se adelantó un ejercicio de estimación del ingreso de las unidades de gasto a partir de información socio-económica de esta fuente, que ilustra una forma de obtener información indirecta sobre esta variable cuando no se disponga de ella<sup>21</sup>.

La evaluación sobre la información relativa a los ingresos consistió en la revisión número de observaciones con información de ingresos (superior a cero) para los distintos conceptos. La base del SISBÉN IV no distingue entre valores omitidos y ausencia de información (valores faltantes), así que se asume que los ceros incluyen ambas situaciones. De esta forma, se considera que una proporción elevada de estos valores expresa una calidad deficiente de la información, con una confiabilidad no aceptable para expresar adecuadamente el ingreso de los individuos y los hogares. Para esta evaluación se recurrió a una tabla de frecuencias relativas de todas las variables referidas a ingreso incluidas dentro de la base. El resultado de este procedimiento se encuentra en el Cuadro 17.

*Cuadro 17. Porcentaje de personas registradas en el SISBÉN IV, sin ingresos, según concepto de ingreso. Bogotá.*

<i>Concepto de ingresos</i>	<i>% sin ingreso</i>
<i>Salarios</i>	78,5
<i>Honorarios</i>	82,0
<i>Ingresos cosecha</i>	100,0
<i>Ingresos pensión</i>	96,9
<i>Ingresos remesas dentro del país</i>	96,3
<i>Ingresos remesas del exterior</i>	99,8
<i>Arriendos</i>	97,8
<i>Familias en Acción</i>	99,4
<i>Colombia mayor</i>	98,1
<i>Sin ingresos (Per cápita)</i>	48,6

*Fuente: Procesamiento propio de las bases de datos del SISBÉN IV.*

Se encontró que la proporción de personas sin ingreso per cápita en el hogar (con cero ingresos) es en el SISBÉN IV de 48,6%. En la EMB 2021 esa proporción es de 4,2%.

<sup>21</sup> Véase el numeral 4.3.5. y el Anexo 2 del documento del producto 3 del proyecto de que forma parte este texto.

Para las variables que miden los distintos tipos de ingresos los valores en cero superan en todos los casos el 40%. Para el caso del salario el porcentaje de personas sin registro es de 78,5% para el de honorario de 82,0% Y sobre el ingreso proveniente de cosechas o ganancias por actividades agrícolas no es reportado por ninguna persona.

Frente al valor de pensiones el porcentaje de personas sin ingresos es de 96,9%, seguido de remesas del país con 96,3%, remesas del exterior 99,8%, y los subsidios no superaron del 2% de los registros en ningún caso. Sumados todas las fuentes de ingreso el reporte Máximo se ubica en 51,4%.

Esta situación da base para concluir que la variable ingreso de la base del SISBÉN IV no es de una calidad aceptable. Tiene un sub-reportaje significativo, que le impide hacer un aporte efectivo a la construcción del RUI.

### 6.3.2. Estimación de los ingresos de las unidades de gasto a partir de variables socioeconómicas del SISBEN IV utilizando la técnica de regresión

Ante las limitaciones de la información directa sobre ingresos se llevó a cabo el ejercicio anunciado de estimar los ingresos de las unidades de gasto a partir de información socioeconómica registrada en la misma encuesta del SISBEN IV. Para realizarlo, se aplicó un modelo de regresión lineal generalizado para calcular el logaritmo del ingreso per cápita a partir de variables socioeconómicas contenidas en la EMB 2021, que también se encontraran en el SISBEN IV. Los coeficientes obtenidos (parámetros) del modelo fueron empleados posteriormente para estimar el ingreso de las unidades de gasto en la base misma del SISBEN.

Se aborda, de esta forma y de una manera práctica, el problema de estimar los ingresos de un amplio conjunto de la población que carece de información directa sobre esta variable en los RA, parte importante del cual es ubicada generalmente en el trabajo informal. La información del SISBEN IV es la única fuente de información de RA a la que se tuvo acceso, que permite hacer este ejercicio de estimación del ingreso para las unidades de gasto que no cuentan con información directa sobre esta variable.

La especificación del modelo con variables comunes a las dos fuentes se expresa de la forma siguiente:

$$\log Y_i = \alpha + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + \beta_4 x_{4i} + \beta_5 x_{5i} + \beta_6 x_{6i} + \beta_7 x_{7i} + \beta_8 x_{8i} + \beta_9 x_{9i} + \beta_{10} x_{10i} + \beta_{11} x_{11i} + \beta_{12} x_{12i} + \beta_{13} x_{13i} + \beta_{14} x_{14i} + \beta_{15} x_{15i} + \beta_{16} x_{16i} + \epsilon_i$$

Las variables utilizadas para aplicar el modelo están reseñadas en el cuadro 18. Y en el cuadro 19 están consignados los coeficientes obtenidos del modelo, con los valores de las pruebas de significancia aplicadas. El coeficiente de determinación ( $R^2$ ) del modelo es de 23,67.

*Cuadro 18. Variables utilizadas en el modelo de regresión utilizado para estimar el ingreso per cápita de las unidades de gasto con la encuesta del SISBÉN IV.*

Variables	Definición	Posibles valores
<b>Número de personas</b>	Número de personas en el hogar	- Una persona - Dos o tres personas - Cuatro o más personas
<b>Jefatura del hogar</b>	Sexo al nacer del jefe de hogar	- Masculino - Femenino
<b>Menores de edad</b>	Al menos una persona menor de edad en el hogar	- Si - No
<b>Personas trabajando</b>	Al menos una persona ocupada	- Si - No
<b>Personas con discapacidad</b>	Al menos una persona con discapacidad en el hogar	- Si - No
<b>Tenencia del hogar</b>	Vivienda ocupada por el hogar propia (totalmente pagada o pagando)	- Si - No
<b>Nivel socioeconómico</b>	Estrato para la tarifa de energía	- Sin estrato - Estrato - Estrato 2 - Estrato 3 - Estrato 4 - Estrato 5 - Estrato 6
<b>Nivel educativo jefe</b>	Nivel educativo del jefe del hogar	- Primaria o menos - Secundaria o media - Técnico o tecnológico - Universidad o más

*Fuente: Cálculos propios obtenidos de la EMB 2021.*

*Cuadro 19. Coeficientes obtenidos del modelo de regresión aplicado a la EMB 2021 para estimar el ingreso per cápita de las unidades de gasto con información del SISBÉN IV-2021*

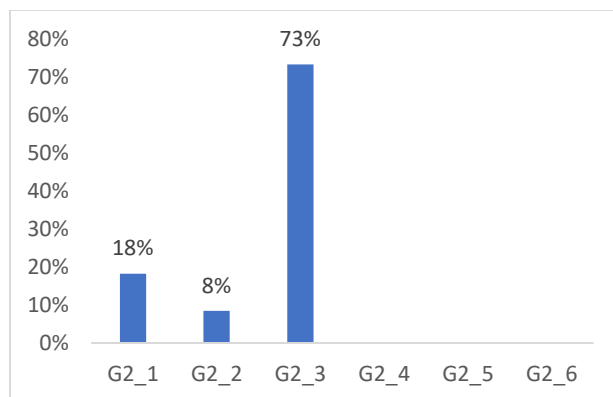
Variables	Estimación	Error estándar	Valor t	Pr(> t )	Significancia
<b>Intercepto</b>	13,32	0,26	51,93	0,00	(***)
<b>Número de personas: 2 o 3 personas</b>	0,22	0,03	7,26	0,00	(**)
<b>Número de personas: 4 o más personas</b>	0,17	0,04	4,21	0,00	(**)
<b>Jefatura del hogar: Mujer</b>	-0,24	0,02	-11,05	0,00	(**)
<b>Menores de edad: No</b>	0,38	0,03	13,54	0,00	(**)
<b>Personas trabajando: No</b>	-2,95	0,02	-120,30	0,00	(***)
<b>Personas con discapacidad: No</b>	0,19	0,02	7,52	0,00	(**)
<b>Tenencia del hogar: No</b>	-0,35	0,02	-15,48	0,00	(**)
<b>Nivel socioeconómico: 1</b>	-0,46	0,25	-1,80	0,07	(.)
<b>Nivel socioeconómico: 2</b>	-0,25	0,25	-1,00	0,32	
<b>Nivel socioeconómico: 3</b>	0,10	0,25	0,42	0,68	
<b>Nivel socioeconómico: 4</b>	0,57	0,25	2,25	0,02	(.)
<b>Nivel socioeconómico: 5</b>	0,97	0,26	3,78	0,00	(**)
<b>Nivel socioeconómico: 6</b>	1,12	0,26	4,32	0,00	(**)
<b>Nivel educativo jefe: Secundaria o media</b>	0,20	0,03	6,74	0,00	(**)
<b>Nivel educativo jefe: Técnico o tecnológico</b>	0,31	0,04	7,95	0,00	(**)
<b>Nivel educativo jefe: Universidad o más</b>	0,76	0,04	21,56	0,00	(**)

*Fuente: Cálculos propios obtenidos de la EMB 2021.*

Cuando se aplican estos coeficientes a la base del SISBÉN IV se calcula el ingreso estimado, a partir del cual se realiza la clasificación de las unidades de gasto en los grupos del modelo 2.

En el Gráfico 44 se observa que, al hacer esta clasificación, en el grupo 3 se concentra el 73% de los hogares, un total de 835.521 hogares, seguido del grupo 1 con 18% de hogares, y del grupo 2 donde se encuentran 8% de los hogares. En los grupos 4, 5 y 6 no se encuentran hogares clasificados.

*Gráfico 44. Distribución porcentual de los hogares del SISBÉN IV, por grupos de ingreso RUI del modelo 2 del RUI-2021*



*Fuente: Cálculos propios obtenidos de la EMB 2021.*

En el grupo 3 del modelo 2 se encuentra la mayor parte de la población registrada en el SISBÉN IV, :aproximadamente 2'7 millones de personas, y en los grupos 1 y 2 se encuentran 410 mil y 436 mil personas respectivamente (Cuadro 20).

*Cuadro 20. Distribución de la población registrada en el SISBÉN IV, por grupos de niveles de ingreso per cápita del modelo 2- 2021.*

Grupos modelo 2	Total Personas
<b>G2_1</b>	410.147
<b>G2_2</b>	436.420
<b>G2_3</b>	2.728.836
<b>Total</b>	3.575.403

*Fuente: Cálculos propios obtenidos de la EMB 2021.*

El ingreso promedio mensual estimado por persona en el grupo 1 es de \$38.006, en el grupo 2 de 2 \$411.479 y en el grupo 3 de \$805.813.



En los gráficos 45 a 49 se hacen comparaciones entre algunas características de los grupos 1, 2, y 3 del modelo 2, obtenidos con la EMB 2021, y los identificados a partir de las estimaciones realizadas con información del SISBÉN IV. Se busca, de esta forma, tener una apreciación aproximada sobre el grado de coincidencia de las poblaciones diferenciadas con estas dos fuentes de información.

En general, se encuentra que no hay coincidencia entre las variables comparadas. Esto da lugar a concluir que los grupos identificados en las dos fuentes, no son los mismos. Se trata de una conclusión que requiere de evaluaciones más detalladas, sustentadas en la consideración de las metodologías y conceptos de las dos encuestas y en el examen de la calidad de sus datos, en lo posible, con confrontaciones sobre el terreno y observación directa de los hogares. En principio, el hecho de que la población encuestada perciba que el acceso a beneficios como los subsidios públicos depende de las respuestas que den a las preguntas de la encuesta SISBÉN, es un factor que incide significativamente en contra de la obtención de información confiable con esta fuente. De otra parte, una herramienta útil en esa evaluación sería la triangulación con información de ingresos y características socioeconómicas provenientes de otras fuentes, que se crucen con base en los números de identificación personal, como la provenientes de catastro distrital, estratificación socioeconómica, programas subsidiados y declaración de renta para impuestos nacionales (DIAN).

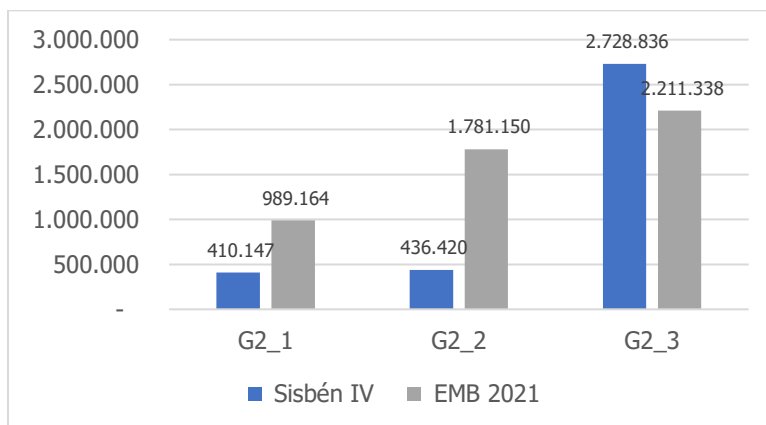
En el gráfico 45 se pueden percibir las no pequeñas diferencias entre la estimación sobre tamaño poblacional de los grupos 1 y 2 del modelo 2, que aparecen considerablemente subvaluados en las estimaciones del SISBÉN, mientras que el 3 se encuentra sobreestimado, en comparación con los datos, más confiables, de la EMB 2021.

Los niveles de los ingresos, que son el aspecto central de la evaluación, difieren en una alta proporción entre los tres grupos. Las estimaciones con el SISBÉN IV están siempre por debajo, en más de 50%, de las de la EMB 2021 (Gráfico 46).

Las comparaciones sobre características demográficas, que corresponden con parámetros básicos para confirmar que se están comparando los mismos grupos (tamaño de los hogares, promedio de edad, composición por sexo) muestran igualmente diferencias apreciables que invalidan la hipótesis de la equivalencia de los grupos comparados (Gráficos 47 a 49).

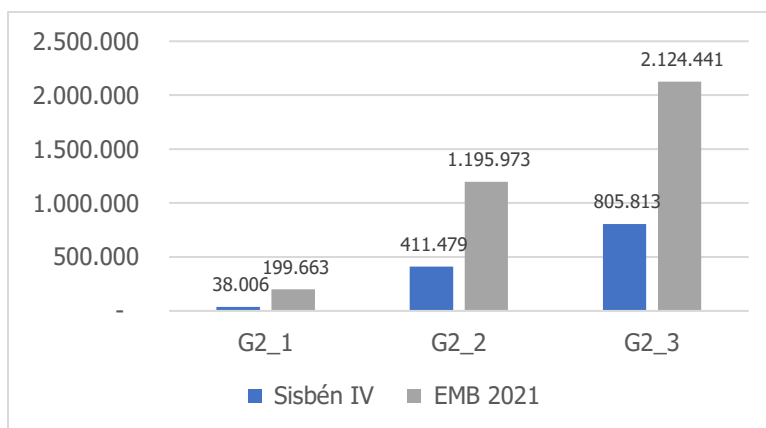
En todo caso, sin embargo, el ejercicio realizado muestra una posible ruta para la estimación del ingreso de las personas y hogares que carecen de información directa sobre esta variable, y para las cuales se puede obtener información socioeconómica que dé una idea de su nivel de vida. Señala una vía que es preciso seguir explorando en la perspectiva del progresivo reemplazo del SISBÉN como instrumento de focalización, por el RUI.

Gráfico 45. Población en los grupos 1, 2 y 3 del modelo 2 del Índice RUI, con la EMB 2021 y con la estimación a partir del SISBÉN IV



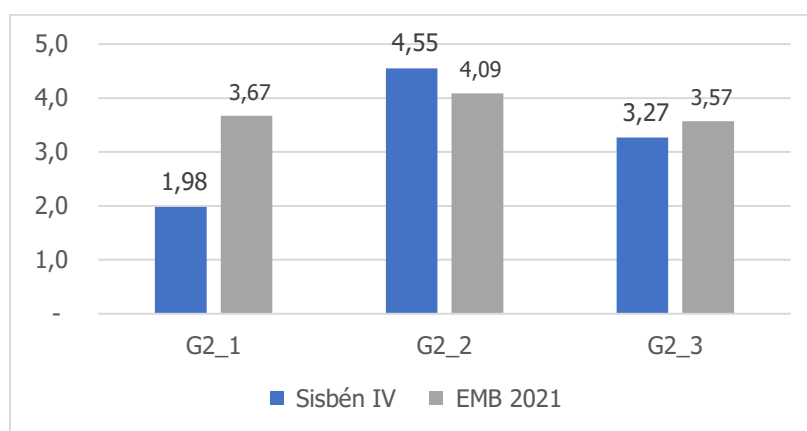
Fuente: procesamiento propio de la EMB 2021 y de las bases del SISBÉN IV.

Gráfico 46. Ingreso promedio por unidad de gasto de los grupos 1, 2 y 3 del modelo 2 del índice RUI según la EMP 2021 y el SISBÉN IV



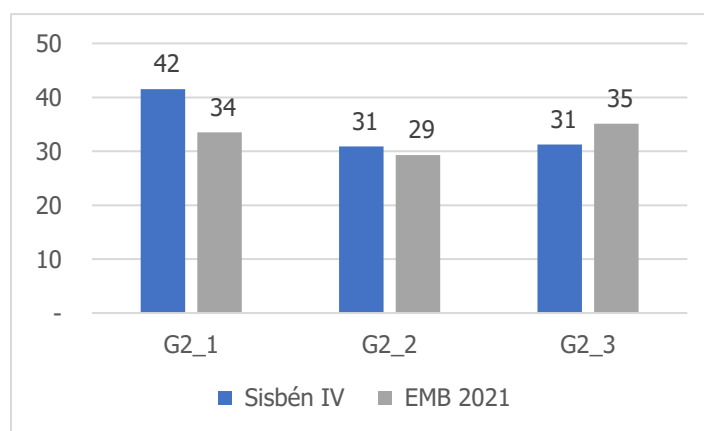
Fuente: procesamiento propio de la EMB 2021 y de las bases del SISBÉN IV.

Gráfico 47. Promedios de personas por hogar por grupos del modelo 2, obtenidos de la EMP 2021 y del SISBÉN IV



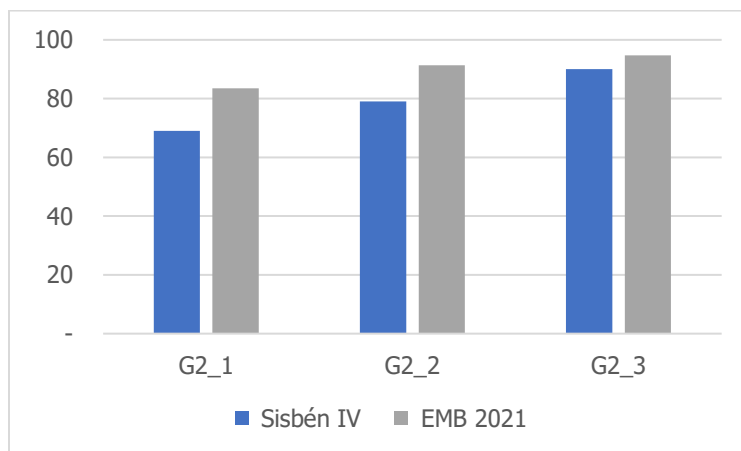
Fuente: procesamiento propio de la EMB 2021 y de las bases del SISBÉN IV.

Gráfico 48. Edad promedio por grupos del modelo 2, obtenidas de la EMP 2021 y del SISBÉN IV



Fuente: procesamiento propio de la EMB 2021 y de las bases del SISBÉN IV.

Gráfico 49. Razón de masculinidad, por grupos del modelo 2, obtenidas de la EMP 2021 y del SISBÉN IV



Fuente: procesamiento propio de la EMB 2021 y de las bases del SISBÉN IV.

## 7. RECOMENDACIONES DE POLÍTICA PÚBLICA.

En el marco de los productos de la consultoría se definió el Registro Universal de Ingresos (RUI) como un Registro Administrativo (RA) que incorpora información de diversas fuentes de información también con características de registros administrativos<sup>22</sup>. Su configuración y forma de operar forma parte del Sistema Estadístico Nacional. El DANE, como ente rector del ese Sistema, ha desarrollado una serie de lineamientos que deben cumplir los RA de tal modo que garanticen la transmisión de la información, así como la interoperabilidad de las bases de datos (arquitectura, plataforma estándares y lineamientos). En este sentido, el RUI una vez conformado deberá cumplir con los estándares de calidad estadística definidos por el DANE.

Las experiencias internacionales, cuya revisión se realizó en el Producto 1 de la consultoría<sup>23</sup>, identifican en registros administrativos personas y familias en situación de vulnerabilidad socio-económica, según la declaración de renta anual, lo cual les confiere derechos a subsidios o subvenciones que otorga el Estado, que consisten en medios de pago de diferentes beneficios de transferencias monetarias que se otorgan con el objetivo de mejorar el nivel de ingresos y el acceso a un consumo básico de bienes y servicios.

En otros términos, la variable fundamental en el RUI es la declaración de ingresos de las personas y las familias, la cual permite una valoración base-inicial de la situación socioeconómica. Esta información, en el caso colombiano, se obtiene de la DIAN. A esta se añade la medición de carencias, que en el caso de Italia se mide con el Índice de la Situación Económica Equivalente (ISEE)<sup>24</sup>, en Uruguay mediante el Índice de Carencias

---

<sup>22</sup> Analytica (2023). Diseño conceptual y metodológico para La construcción del registro universal de ingresos (producto 3, versión 2), p. 45. Mimeo. Documento presentado en el marco del Contrato 707 de 2023 suscrito con la Secretaría Distrital de Planeación con el objeto de *Analizar la estratificación como instrumento de focalización junto con otros instrumentos de política pública para determinar su eficiencia*.

En el Sistema Estadístico Nacional (SEN) el Registro Administrativo (RA) se define como un conjunto de datos que contiene la información recogida y conservada por entidades u organizaciones en el cumplimiento de sus funciones o competencias misionales. Ibarra Duarte Maritza y Fonseca Lemus Arlington (2018). Registros Administrativos y operaciones estadísticas de Función Pública. Departamento Administrativo de la Función Pública - Oficina Asesora de Planeación. Recuperado el 04 de enero de 2024 de [https://www.funcionpublica.gov.co/VisualSIE/faces/javax.faces.resource/docs/Documento\\_operaciones\\_estadisticas\\_y\\_registros\\_administrativos.pdf](https://www.funcionpublica.gov.co/VisualSIE/faces/javax.faces.resource/docs/Documento_operaciones_estadisticas_y_registros_administrativos.pdf). En el inventario de registros administrativos del DANE se señala que Colombia cuenta con 393 registros administrativos generados por 144 entidades de carácter público, privado y mixto. DANE (2024). Sistema de Identificación y Caracterización de Oferta y Demanda Estadística del Sen. Inventario de Registros Administrativos. Recuperado el 04 de enero de 2024 de [https://inventariosen.dane.gov.co/rrea/consulta\\_rrea](https://inventariosen.dane.gov.co/rrea/consulta_rrea)

<sup>23</sup> Analytica (2023). Descripción del cobro de servicios públicos domiciliarios y su focalización. Registro universal de ingresos, declaración y autodeclaración de renta (países de la OCDE y de América Latina). Mimeo. Documento presentado en el marco del Contrato 707 de 2023 suscrito con la Secretaría Distrital de Planeación con el objeto de *Analizar la estratificación como instrumento de focalización junto con otros instrumentos de política pública para determinar su eficiencia*.

<sup>24</sup> Documento que certifica y mide la situación económica de personas y familias, y contiene los datos personales (nombre, apellido, fecha de nacimiento, código fiscal, dirección de residencia), de ingresos (de trabajo dependiente, de trabajo autónomo y otros) y de patrimonio (vivienda, propiedades, saldos de cuentas

Críticas (ICC)<sup>25</sup>, en Perú a través del SISFOH (sistema de focalización de hogares)<sup>26</sup>, mecanismo que permite identificar personas y familias en situación de pobreza para acceder a los programas sociales, mediante clasificación socioeconómica de las familias (CSE) como pobre, pobre extremo o no pobre.

En todos los casos se contrasta la declaración de renta y la situación de empleo o desempleo de las personas, esta última a través de los ministerios de trabajo o similares. La información de los países de la OCDE muestra los avances que han tenido en materia de focalización de los recursos estatales para atender necesidades de la población en situación de vulnerabilidad. La experiencia de Dinamarca con la Tarjeta Azul es el caso más paradigmáticos en tanto relaciona aspectos tributarios y aspectos sociales de una manera en la que el mercado de trabajo, la seguridad social y los establecimientos educativos se han convertido en el medio idóneo de la información.

Experiencias similares a la de Dinamarca se ha observado en Italia, Argentina, Uruguay y Chile, con diferencias, desde luego, pero en sentido amplio apuntan a los objetivos de configurar un Registro Universal con interconexiones con bases de datos claves que permiten identificar y verificar las necesidades reales de los ciudadanos usuarios que solicitan apoyos estatales.

En Italia, la condición para el acceso a cualquier programa social gubernamental es el registro de condiciones de vida de la persona y del grupo familiar. Lo mismo se exige en Argentina, Uruguay y Chile, en este último a través del Registro Social de Hogares, que es la experiencia que en Colombia ha comenzado a desarrollarse con mayor vigor en el marco del actual Plan de Desarrollo.

La característica más notable y coincidente entre los países de la OCDE y los de América Latina incluidos en el estudio con mayor avance en el Registro Universal de Ingresos es la de que son las personas quienes deben solicitar inscribirse en el registro de su país para acceder a los apoyos estatales. La otra características coincidente es la de los temas que tanto en los países de la OCDE como los de América Latina se tienen en cuenta grupos poblacionales que son prioridad en la atención estatal desde la perspectiva de los derechos, entre niños, niñas y adolescentes, personas en discapacidad, número de hijos, adulto mayor, situación de vulnerabilidad.

Según las experiencias internacionales, las unidades de análisis en sistemas de información basados en registros administrativos son el individuo y la familia<sup>27</sup>. Estas

---

bancarias y otros). Gobierno de Italia (2023). Sistema de información sobre el Indicador de la Situación Económica Equivalente (ISEE). Recuperado el 23 de octubre de 2023 de <https://sociale.comune.fi.it/system/files/2022-08/CORRETTO%20SPAGNOLO%20ISSE%2030%20GEN%2022%20WEB.pdf>

<sup>25</sup> Gobierno de Uruguay – Ministerio de Desarrollo Social (2023). Tarjeta Uruguay Social -TUS. Recuperado el 23 de noviembre de 2023 de <https://www.gub.uy/ministerio-desarrollo-social/node/9812>.

<sup>26</sup> Gobierno de Perú - Ministerio de Desarrollo e Inclusión Social (2023). Sistema de Focalización de Hogares (SISFOH). Recuperado el 28 de octubre de 2023 de <https://www.gob.pe/458-sistema-de-focalizacion-de-hogares-sisfoh-clasificacion-socioeconomica-para-poblaciones-amazonicas>

<sup>27</sup> Según el DANE, en la variable individuo se observan las características sociodemográficas de la persona (género, edad, entre otras) y económicas (ingreso, educación, entre otras). En el caso de encuestas y censos,

unidades deben corresponder al RUI, en la perspectiva del análisis social que debe soportar y en su papel de apoyar una forma eficaz de focalización.

En la perspectiva de las experiencias internacionales y los alcances del Registro Social de Hogares (RSH) en el propósito de construir el Registro Universal de Ingresos desde el actual Plan Nacional de Desarrollo<sup>28</sup>, dan las pautas para orientaciones de política en la creación del RUI, en una secuencia jerárquica de integración de registros administrativos, que se pueden denominar niveles<sup>29</sup>, así:

- Nivel 1: Información de declaración de renta (DIAN) y de aportes a la seguridad social y al sistema de pensiones en los regímenes contributivos y subsidiado (información de PILA, del Ministerio de Salud y Protección Social, y de subsidios a través de ADRES). Para el desarrollo del RUI en el Distrito Capital de Bogotá, la información se podrá obtener de la DIAN a través de la Secretaría Distrital de Hacienda (SDH) y la información de PILA y de ADRES a través de la Secretaría Distrital de Salud (SDS). Para ese propósito, en la vía del Plan Nacional de Desarrollo, el Distrito Capital deberá establecer un convenio interadministrativo con el DNP, como un medio expedito para el acceso a la información.
- Nivel 2: Indicadores de carencias o de clasificación socioeconómica de personas y familias (pobreza extrema, pobreza, no pobreza). En este nivel de información se identifican grupos poblacionales vulnerables de especial atención por parte del Estado, en la perspectiva de los derechos.

Esta información amplía el espectro de variables necesarias para la identificación de situaciones de vulnerabilidad de personas, familias y grupos de especial protección en la perspectiva de los derechos. Inicialmente la información se obtiene de encuestas -como la EMB utilizada en el ejercicio de esta consultoría, y el SISBEN-. Desde luego, la construcción y consolidación del RUI, sobre la base de integración de bases de datos de registros administrativos, conducirá a desestimar en el transcurso del tiempo este tipo de información.

Uno de los aspectos esenciales identificado en los programas sociales de los países de la OCDE es el de la tarifa de los servicios públicos domiciliarios de energía (electricidad y gas) y agua potable y saneamiento básico. Las personas y las familias que se encuentran en situación de vulnerabilidad (que en el caso colombiano se pueden identificar según los

---

la definición de hogar, desde el DANE, se refiere a un individuo o grupo de individuos, parientes o no, que ocupan total o parcialmente una vivienda, atienden a necesidades básicas con un presupuesto común y, generalmente, comparten comidas. Se tiene información sobre el número de individuos que pertenecen al hogar y variables agregadas (sumas o promedios) para el conjunto del hogar. Con los registros administrativos se obviaría la definición de hogar y se centraría en la definición de familia, como ocurre en los países de la OCDE.

<sup>28</sup> Plan Nacional de Desarrollo 2022-2026 Colombia Potencia Mundial de la Vida (Congreso de la República, 2023). El Artículo 70 de este plan se refiere a la creación del RUI, cuya administración queda a cargo del Departamento Nacional de Planeación.

<sup>29</sup> En la experiencia de Uruguay, los niveles permiten la identificación de personas y familias en situación de vulnerabilidad extrema (nivel 1) y grupos especiales vulnerables (LGTBI, niños, niñas y adolescentes -NNA-, personas de tercera edad, entre otros (nivel 2).

niveles 1 y 2 mencionados), podrán solicitar apoyos monetarios estatales. Para el efecto, la dinámica de los precios de mercado de estos servicios permite de manera general identificar los impactos que pueden tener en la capacidad de gasto de personas y familias, a partir de los cuales se podrán considerar medidas orientadas a mitigar tales impactos, mediante subsidios temporales. Esta posibilidad supone un RUI consolidado y la desaparición de la estratificación como mecanismo para la asignación de subsidios.

## 7.1 Recomendaciones de política sobre Habeas Data

Conforme a lo previsto en el artículo 1 de la Constitución Política (Constitución Política de la República de Colombia, 1991) *“Colombia es un Estado social de derecho, organizado en forma de República unitaria, descentralizada, con autonomía de sus entidades territoriales, democrática, participativa y pluralista, fundada en el respeto de la dignidad humana, en el trabajo y la solidaridad de las personas que la integran y en la prevalencia del interés general”*.

En virtud del principio de solidaridad, el Estado Colombiano se ha propuesto apoyar a la población formulando políticas públicas tendientes a mejorar progresivamente su calidad de vida y eliminar la pobreza. Varios de esos aspectos se pueden observar a nivel nacional con programas como: Adjudicación de baldíos, Atención Integral a la Primera Infancia, Protección (Hogar gestor y otros similares), Matricula Cero, Subsidio Gas Licuado de Petróleo – GLP, Régimen Subsidiado de Salud, Casa Digna Vida Digna, Mi casa Ya, Colombia Mayor, Programa de Alimentación Escolar PAE, Atención Humanitaria, Apoyo Sostenimiento, entre otros.

En el Distrito Capital se identificaron diecisiete (17) programas que involucran la estratificación social, los cuales se encuentran descritos en el producto No. 2 de la presente consultoría<sup>30</sup>, cuyo propósito es otorgar diversos beneficios a la población (Cuadro 21).

*Cuadro 21. Programas del orden distrital que involucran estratificación social*

Entidad	Programa
Secretaría Distrital de Hacienda	Sistema Tributario Distrital Impuesto de delimitación urbana en el Distrito Capital
Secretaría de Educación del Distrito	Fondo Solidario de Educación
Secretaría Distrital de Integración Social	Comedores Comunitarios – Cocinas Populares
Secretaría Distrital de Hábitat	Programa Mejoramiento de Vivienda
Instituto para la Economía Social IPES	Apoyo y acompañamiento para la inclusión financiera de los emprendimientos.
Instituto Distrital de Patrimonio Cultural IDPC	Incentivo de equiparación a estrato uno (1) para el cobro de tarifas de servicios públicos en los bienes de interés cultural
Caja de Vivienda Popular CVP	Asistencia técnica para la obtención de licencias de construcción y/o actos de reconocimiento Programa de reasentamiento de familias por encontrarse en condiciones de alto riesgo no mitigable en el Distrito Capital
Empresa Metro de Bogotá S.A	Plan de Reasentamiento para la Primera Línea del Metro de Bogotá

<sup>30</sup> Documento soporte del inventario de los programas distritales que requieren estratificación social.



Entidad	Programa
Empresa de Acueducto y Alcantarillado de Bogotá EAAB-ESP	Mínimo vital Acuerdos de pago
Instituto Distrital de Protección y Bienestar Animal	Esterilizar salva
	Ciudadanos de cuatro patas
	Tarifa estudiantil SITP
SITP	Costo recolección de escombros
IDU	Cobro por valorización
Universidad Distrital Francisco José de Caldas	Valor matrícula

*Fuente: Elaboración Propia*

Lo anterior implica que cada una de las entidades que tienen a cargo los programas antes referidos, cuentan con información relevante para determinar la población de potenciales beneficiarios a programas sociales, surgiendo por tanto el cuestionamiento frente al cruce de información entre entidades a propósito de consolidar una base de datos que permita identificar con mayor diligencia a la población vulnerable.

Para abordar el cuestionamiento planteado, es necesario referir que el artículo 15 que prevé:

*“todas las personas tienen derecho a su intimidad personal y familiar y a su buen nombre, y el Estado debe respetarlos y hacerlos respetar. De igual modo, tienen derecho a conocer, actualizar y rectificar las informaciones que se hayan recogido sobre ellas en bancos de datos y en archivos de entidades públicas y privadas. En la recolección, tratamiento y circulación de datos se respetarán la libertad y demás garantías consagradas en la Constitución.*

*La correspondencia y demás formas de comunicación privada son inviolables. Sólo pueden ser interceptadas o registradas mediante orden judicial, en los casos y con las formalidades que establezca la ley.*

*Para efectos tributarios o judiciales y para los casos de inspección, vigilancia e intervención del Estado podrá exigirse la presentación de libros de contabilidad y demás documentos privados, en los términos que señale la ley.”*

Frente al derecho fundamental al habeas data contenido en el artículo en mención, la Corte Constitucional en la Sentencia T-238-18, expresó:

*“El derecho fundamental al hábeas data se encuentra consagrado en el artículo 15 Superior que dispone que todas las personas tienen derecho a la intimidad personal, al buen nombre, a conocer, actualizar y rectificar la información que se haya recogido sobre ellas en los diferentes bancos de datos y en los archivos de entidades públicas y privadas. Adicionalmente, establece la obligación que tiene el Estado de hacer respetar tales derechos. Asimismo, de conformidad con el artículo 152 de la Constitución Política, corresponde al Congreso de la República regular los derechos fundamentales de las personas, los procedimientos y recursos para su protección a través de la expedición de leyes estatutarias.”*

El derecho en mención se encuentra desarrollado en la Ley 1266 de 2008, la cual se aplica frente a todos los datos de información personal registrados en un banco de datos, independiente de la naturaleza de la entidad que los administre (pública o privada). La

norma en mención es absolutamente relevante en tanto clasifica y define los tipos de datos, así mismo alude los principios para la administración de aquellos.

En relación con la clasificación de los datos, se resalta que el artículo 3 de la ley en referencia alude lo siguiente:

*“Artículo 3. Definiciones. Para los efectos de la presente ley, se entiende por: (...)*

*e) Dato personal. Es cualquier pieza de información vinculada a una o varias personas determinadas o determinables o que puedan asociarse con una persona natural o jurídica. Los datos impersonales no se sujetan al régimen de protección de datos de la presente ley. Cuando en la presente ley se haga referencia a un dato, se presume que se trata de uso personal. Los datos personales pueden ser públicos, semiprivados o privados;*

*f) Dato público. Es el dato calificado como tal según los mandatos de la ley o de la Constitución Política y todos aquellos que no sean semiprivados o privados, de conformidad con la presente ley. Son públicos, entre otros, los datos contenidos en documentos públicos, sentencias judiciales debidamente ejecutoriadas que no estén sometidos a reserva y los relativos al estado civil de las personas;*

*g) Dato semiprivado. Es semiprivado el dato que no tiene naturaleza íntima, reservada, ni pública y cuyo conocimiento o divulgación puede interesar no sólo a su titular sino a cierto sector o grupo de personas o a la sociedad en general, como el dato financiero y crediticio de actividad comercial o de servicios a que se refiere el Título IV de la presente ley.*

*h) Dato privado. Es el dato que por su naturaleza íntima o reservada sólo es relevante para el titular”.*

Así la norma, los datos pueden ser personales, públicos, semiprivados y privados, correspondiendo al responsable de su tratamiento garantizar su protección.

La Ley 1581 de 2012 (Ley 1581 , 2012) *“tiene por objeto desarrollar el derecho constitucional que tienen todas las personas a conocer, actualizar y rectificar las informaciones que se hayan recogido sobre ellas en bases de datos o archivos, y los demás derechos, libertades y garantías constitucionales a que se refiere el artículo 15 de la Constitución Política; así como el derecho a la información consagrado en el artículo 20 de la misma”,* trae consigo una categoría especial de datos, esto es, datos sensibles, los cuales define así:

*“Artículo 5. Datos sensibles. Para los propósitos de la presente ley, se entiende por datos sensibles aquellos que afectan la intimidad del Titular o cuyo uso indebido puede generar su discriminación, tales como aquellos que revelen el origen racial o étnico, la orientación política, las convicciones religiosas o filosóficas, la pertenencia a sindicatos, organizaciones sociales, de derechos humanos o que promueva intereses de cualquier partido político o que garanticen los derechos y garantías de partidos políticos de oposición así como los datos relativos a la salud, a la vida sexual y los datos biométricos”.*

Ahora bien, en su capítulo IV de derechos y condiciones de legalidad para el tratamiento de datos, la norma en el artículo 10 enlista los casos en los que no es necesaria la autorización del titular para el tratamiento de los datos, entre ellas cuando se trate de:

*“a) Información requerida por una entidad pública o administrativa en ejercicio de sus funciones legales o por orden judicial”*

La norma en cita se armoniza con lo previsto en el artículo 14 de la Ley 962 de 2005 (Ley 962 , 2005) que modificó el artículo 16 del Decreto-ley 2150 de 1995, así:

*“ARTÍCULO 16. Solicitud oficiosa por parte de las entidades públicas. Cuando las entidades de la Administración Pública requieran comprobar la existencia de alguna circunstancia necesaria para la solución de un procedimiento o petición de los particulares, que obre en otra entidad pública, procederán a solicitar a la entidad el envío de dicha información. En tal caso, la carga de la prueba no corresponderá al usuario.*

***Será permitido el intercambio de información entre distintas entidades oficiales, en aplicación del principio de colaboración.***

*El envío de la información por fax o cualquier otro medio de transmisión electrónica, proveniente de una entidad pública, prestará mérito suficiente y servirá de prueba en la actuación de que se trate siempre y cuando se encuentre debidamente certificado digitalmente por la entidad que lo expide y haya sido solicitado por el funcionario superior de aquel a quien se atribuya el trámite.*

*Cuando una entidad pública requiera información de otra entidad de la Administración Pública, esta dará prioridad a la atención de dichas peticiones, debiendo resolverlas en un término no mayor de diez (10) días, para lo cual deben establecer sistemas telemáticos compatibles que permitan integrar y compartir información de uso frecuente por otras autoridades”. (Negrilla fuera de texto original)*

El principio de colaboración frente al intercambio de información no es absoluto ni ilimitado so pena de vulnerar el derecho al habeas data. La Sala de Consulta y Servicio Civil del Consejo de Estado dentro del radicado número: 11001-03-06-000-2020-00234-00(2458), analizó esta circunstancia, determinando que *“la información que puede entregarse a una autoridad administrativa sin la autorización del titular de los datos personales es aquella que corresponde a datos públicos y semiprivados. En este último caso, se requiere que la información sea necesaria para el cumplimiento de sus funciones. En lo que respecta a los datos privados y sensibles, a ellos no puede acceder la autoridad sin la autorización del titular”*. Lo anterior, resulta apropiado bajo el entendido que:

- (i) Por su naturaleza y en atención a lo dispuesto en el artículo 74 de la Constitución Política, a los datos considerados como públicos puede acceder cualquier persona y, por consiguiente, pueden ser entregados a la autoridad que los requiera.
- (ii) En cuanto a los datos semiprivados la Corte Constitucional en la Sentencia T-058/15 concluyó que este tipo de información está sometida a un grado mínimo de limitación, por lo que puede ser obtenida y ofrecida por una autoridad administrativa en cumplimiento de sus funciones.

- (iii) En la misma Sentencia T-058/15, la Corte Constitucional refirió que *“la información privada puede tener contenidos personales o no y por encontrarse en un ámbito puramente privado, sólo puede ser obtenida y ofrecida por orden de autoridad judicial en el cumplimiento de sus funciones”*, lo cual conduce a determinar que la misma no puede ser entregada únicamente al amparo del principio de colaboración.
- (iv) En cuanto a los datos sensibles, el artículo 6 de la Ley 1581 de 2012 refiere que es necesaria la autorización del titular del dato para su tratamiento.

De acuerdo con lo anterior, se determina que, es posible en virtud del principio de colaboración requerir un intercambio de información entre entidades del Distrito Capital a propósito de construir una base de datos de potenciales beneficiarios de subsidios que requieran estratificación socioeconómica, ello teniendo en consideración que la solicitud se realizaría con fundamento en las funciones propias de la Secretaría Distrital de Planeación y teniendo en consideración la importancia de lo que significa transformar las condiciones de acceso a la población a programas que permitan mejorar las condiciones de vida eliminando la pobreza.

Aunado a lo anterior, es viable se incorpore en el Plan de Desarrollo del Distrito Capital del gobierno que se inicia en este 2024 un artículo en virtud del cual exista un compromiso de las entidades distritales frente al reporte de la información que sea necesaria para consolidar los datos que caractericen la población vulnerable.

## 8. CONCLUSIONES Y RECOMENDACIONES

En la introducción de este documento señalamos que los resultados de la consultoría estaban condicionados a la disponibilidad y utilización de los RA. Así como en las experiencias internacionales de países de la OCDE y de América Latina consultadas se pudo establecer que el certificado de identificación ciudadana y/o certificado de identificación tributaria es la llave para la articulación de bases de datos basadas en registros administrativos, en el ejercicio de esta consultoría no se logró disponer de esta variable. No fue posible obtener información de la DIAN y las bases de datos obtenidas estaban anonimizadas.

En la información de PILA se pudo establecer la distinción entre cotizantes asalariados e independientes, como una aproximación desde la perspectiva de los ingresos laborales, que son una parte de los ingresos primarios corrientes. Para la identificación de la situación de la población general de Bogotá en materia de vinculación al sistema de seguridad social en salud y pensiones, era necesaria además la información bajo el dominio de la ADRES, referida al régimen subsidiado. Sin ella, el análisis que se logró presentar en este Producto 4 es parcial.

Otra de las formas de aproximación al relacionamiento de bases de datos fue a partir del análisis de la información que se obtiene de la EMB 2021 y del SISBÉN IV. De entrada, relacionar un registro administrativo como PILA, con información de encuestas como el del EMB, se debe entender como un ejercicio de aproximación. En la consultoría se propuso como uno de los modelos del índice del RUI el que distingue en la escala los dos primeros niveles asociados a la pobreza extrema y a la pobreza respectivamente y otros cuatro niveles según agrupación de ingresos, para un total de seis, con fines comparativos con la estratificación actual<sup>31</sup>. Así, desde la perspectiva de la información de la EMB, el ejercicio de relacionamiento consistió en establecer la distinción de cotizantes del sistema de seguridad social: asalariados e independientes, como una forma de establecer una comparación aproximativa con la información de PILA. Y de igual modo con la información del SISBÉN IV.

En el Cuadro 20 de este documento se expuso información de PILA, EMB y SISBÉN IV. Hemos señalado las diferencias encontradas entre una base y otra -por ejemplo el número de trabajadores de PILA y el número de trabajadores en la EMB: en cotizantes trabajadores independientes la diferencia entre los segmentos comparados es inferior al 10% y cuanto al número de asalariados en el régimen subsidiados la diferencia entre la PILA y la EMB es de 72%-. Esta situación lleva a concluir que posiblemente hay duplicación en el conteo de un número apreciable de trabajadores en la PILA o el periodo de cotización no coincide con el de trabajo, ya que no es probable que en la encuesta haya un subregistro de trabajadores en la magnitud de lo que lo estaría mostrando su comparación con PILA. Para controlar

---

<sup>31</sup> En algunos países, cuyas experiencias de registros universales fueron consultadas, se establecen distintos niveles de agrupación de ingresos: en países europeos, por ejemplo, ingresos altos, medios y bajos y en países de América Latina, algunas desagregaciones de esos tres niveles.

esas duplicaciones sería necesario contar con las claves de identificación completas de los trabajadores, de las cuales no se dispuso. Y hacer igualmente una revisión de la compatibilidad de las categorías de clasificación de cada una de las dos fuentes.

En cuanto a la información sobre el Ingreso Base de Cotización (IBC) se evidencian discrepancias entre las dos fuentes en cuanto al promedio de las remuneraciones. En algunos grupos las diferencias son grandes. En los afiliados a los regímenes espaciales son de más del 200%. En los empleados domésticos de 21.4% y en las categorías de asalariados e independientes se sitúan en niveles más moderados de -4.2 y 1.8% respectivamente.

En la perspectiva de ir consolidando los pilares básicos del RUI convendría ir regularizando la producción de la información provista por ese RA y ajustando las categorías y forma de procesamiento para logra un registro adecuado y compatible de esa información, con la participación del DANE y el DNP.

Este ejercicio ilustra la compleja tarea que significa la adecuación de los RA, para su integración coherente en el RUI. Y señala cómo los resultados obtenidos impiden la utilización de la PILA, en su conformación actual disponible, para obtener un estimado confiable de los ingresos laborales de los cotizantes al Sistema de Seguridad Social en Salud.

Lo ideal, desde luego, continúa siendo la aspiración de relacionar mediante la identificación de cada individuo y de cada miembro del grupo familiar sus condiciones en materia de ingresos y de características socioeconómicas, para establecer la situación real de vulnerabilidad o no y de posible beneficiarios a los recursos del Estado a través de los diversos programas sociales institucionales.

Entendiendo que para la materialización de una política pública para la aplicación del instrumento de declaración universal, es necesaria la implementación de varias fases, las cuales van desde una fase preparatoria, Agenda pública, Formulación, Implementación, Seguimiento y Evaluación, lo anterior de conformidad con el instrumento de formulación de políticas públicas establecido para el Distrito Capital y formulado por la Secretaría Distrital de Planeación (Alcaldía Mayor de Bogotá, 2017).

En este sentido al considerar el contenido de dicho instrumento, sus distintas fases, los costos que implicaría la formulación efectiva de la política pública, no se estiman necesario teniendo en cuenta el resultado de las actividades desarrolladas en el marco de la consultoría, para el caso lo potencialmente efectos es incorporar un artículo en el plan de desarrollo que se está formulando actualmente, en el que se indique de un lado el deber que recaerá en la Secretaría Distrital de Planeación consolidar la información que se requiere para sentar las bases de la conformación del Registro Universal del Ingresos en el Distrito Capital de modo tal que se puedan refocalizar subsidios y programas en beneficio de la población que lo requiera. Para dicha materialización las entidades del distrito capital que manejen bases de datos de interés para dicho propósito tendrán el deber de remitir la información que se requiera.

Así las cosas, se recaba que en el proceso de construir el RUI para el Distrito Capital, la recomendación de incorporar un artículo similar al artículo 70 del Plan Nacional de



Desarrollo en el Plan de Desarrollo Distrital, es una oportunidad valiosa, en tanto no sólo va en armonía con los propósitos de la Ley a nivel nacional, sino también por el carácter de Distrito que tiene Bogotá y puede contribuir con efecto demostración al desarrollo de este Registro Universal de Ingresos y a su utilización con eficiencia y efectividad en la focalización del gasto público social.

## RECOMENDACIONES DE POLÍTICA

De las experiencias internacionales y los alcances del Registro Social de Hogares (RSH) en el propósito de construir el Registro Universal de Ingresos desde el actual Plan Nacional de Desarrollo, dan las pautas para orientaciones de política en la creación del RUI, en una secuencia jerárquica de integración de registros administrativos, que se pueden denominar niveles, así:

- Nivel 1: Información de declaración de renta (DIAN) y de aportes a la seguridad social y al sistema de pensiones en los regímenes contributivos y subsidiado (información de PILA, del Ministerio de Salud y Protección Social, y de subsidios a través de ADRES). Para el desarrollo del RUI en el Distrito Capital de Bogotá, la información se podrá obtener de la DIAN a través de la Secretaría Distrital de Hacienda (SDH) y la información de PILA y de ADRES a través de la Secretaría Distrital de Salud (SDS). Para ese propósito, en la vía del Plan Nacional de Desarrollo, el Distrito Capital deberá establecer un convenio interadministrativo con el DNP, como un medio expedito para el acceso a la información.
- Nivel 2: Indicadores de carencias o de clasificación socioeconómica de personas y familias (pobreza extrema, pobreza, no pobreza). En este nivel de información se identifican grupos poblacionales vulnerables de especial atención por parte del Estado, en la perspectiva de los derechos.

Es posible, en virtud del principio de colaboración, requerir un intercambio de información entre entidades del Distrito Capital, con el propósito de construir una base de datos de potenciales beneficiarios de subsidios que requieran estratificación socioeconómica, ello teniendo en consideración que la solicitud se realizaría con fundamento en las funciones propias de la Secretaría Distrital de Planeación y teniendo en consideración la importancia de lo que significa transformar las condiciones de acceso a la población a programas que permitan mejorar las condiciones de vida eliminando la pobreza.

Aunado a lo anterior, es viable un proceso de colaboración entre la Administración Distrital y el Departamento Nacional de Planeación en la perspectiva de construcción del RUI. En este estudio se tiene un aporte importante frente a ese objetivo, en tanto los resultados que se presentan incluye la metodología del índice, tarea que el DNP ha previsto desarrollar en el 2025, una vez consolide el Registro Social de Hogares.

## REFERENCIAS

- Analytica (2023). Diseño conceptual y metodológico para La construcción del registro universal de ingresos. Producto 3, versión 2, p. 45. Mimeo. Documento presentado en el marco del Contrato 707 de 2023 suscrito con la Secretaría Distrital de Planeación con el objeto de Analizar la estratificación como instrumento de focalización junto con otros instrumentos de política pública para determinar su eficiencia.
- Analytica (2023). Descripción del cobro de servicios públicos domiciliarios y su focalización. Registro universal de ingresos, declaración y autodeclaración de renta (países de la OCDE y de América Latina). Mimeo. Documento presentado en el marco del Contrato 707 de 2023 suscrito con la Secretaría Distrital de Planeación con el objeto de Analizar la estratificación como instrumento de focalización junto con otros instrumentos de política pública para determinar su eficiencia.
- Congreso de la República de Colombia (1991). Constitución Política de la República de Colombia. Bogotá. <http://secretariassenado.gov.co/index.php/constitucion-politica>.
- Congreso de la República (2001). Ley 715 de 2001, Por la cual se dictan normas orgánicas en materia de recursos y competencias de conformidad con los artículos 151, 288, 356 y 357 (Acto Legislativo 01 de 2001) de la Constitución Política y se dictan otras disposiciones para organizar la prestación de los servicios de educación y salud. [http://www.secretariassenado.gov.co/senado/basedoc/ley\\_0715\\_2001.html](http://www.secretariassenado.gov.co/senado/basedoc/ley_0715_2001.html).
- Congreso de la República (2005). Ley 962 de 2005, Por la cual se dictan disposiciones sobre racionalización de trámites y procedimientos administrativos de los organismos y entidades del Estado y de los particulares que ejercen funciones públicas o prestan servicios públicos. Bogotá. <https://www.funcionpublica.gov.co/eva/gestornormativo/norma.php?i=17004>
- Congreso de la República (2012). Ley Estatutaria 1581 de 2012, por la cual se dictan disposiciones generales para la protección de datos personales. Bogotá. <https://www.funcionpublica.gov.co/eva/gestornormativo/norma.php?i=49981>
- Congreso de la República de Colombia (2023). Ley 2294 del 19 de mayo de 2023, "Por el Cual se expide el Plan Nacional de Desarrollo 2022- 2026 Colombia Potencia Mundial de la Vida". Bogotá. [https://colaboracion.dnp.gov.co/CDT/portalDNP/PND-2023/Ley\\_2294\\_del\\_19\\_de\\_mayo\\_de\\_2023.pdf](https://colaboracion.dnp.gov.co/CDT/portalDNP/PND-2023/Ley_2294_del_19_de_mayo_de_2023.pdf)
- Corte Constitucional – República de Colombia (2015). Sentencia T-058/15. Derecho al Habeas Data. <https://www.corteconstitucional.gov.co/relatoria/2015/T-058-15.htm>
- DANE (2022). Encuesta Multipropósito Bogotá-Cundinamarca 2021. Información actualizada el 01 de julio de 2022. <https://www.dane.gov.co/index.php/estadisticas-por-tema/pobreza-y-condiciones-de-vida/encuesta-multiproposito>.

- DANE (2024). Sistema de Identificación y Caracterización de Oferta y Demanda Estadística del Sen. Inventario de Registros Administrativos. [https://inventariosen.dane.gov.co/rraa/consulta\\_rraa](https://inventariosen.dane.gov.co/rraa/consulta_rraa).
- Congreso de la República (2023). Plan Nacional de Desarrollo 2022-2026 Colombia Potencia Mundial de la Vida. <https://www.dnp.gov.co/plan-nacional-desarrollo/pnd-2022-2026>.
- Dirección de Impuestos y Aduanas Nacionales -DIAN- (2024). [https://www.dian.gov.co/atencionciudadano/formulariosinstructivos/Formularios/2024/Formulario\\_210\\_2024.pdf](https://www.dian.gov.co/atencionciudadano/formulariosinstructivos/Formularios/2024/Formulario_210_2024.pdf)
- Gobierno de Italia (2023). Sistema de información sobre el Indicador de la Situación Económica Equivalente (ISEE). <https://sociale.comune.fi.it/system/files/2022-08/CORRETTO%20SPAGNOLO%20ISSE%2030%20GEN%2022%20WEB.pdf>
- Gobierno de Uruguay – Ministerio de Desarrollo Social (2023). Tarjeta Uruguay Social - TUS. <https://www.gub.uy/ministerio-desarrollo-social/node/9812>.
- Gobierno de Perú - Ministerio de Desarrollo e Inclusión Social (2023). Sistema de Focalización de Hogares (SISFOH). <https://www.gob.pe/458-sistema-de-focalizacion-de-hogares-sisfoh-clasificacion-socioeconomica-para-poblaciones-amazonicas>.
- Ibarra Duarte Maritza y Fonseca Lemus Arlington (2018). Registros Administrativos y operaciones estadísticas de Función Pública. Departamento Administrativo de la Función Pública - Oficina Asesora de Planeación. [https://www.funcionpublica.gov.co/VisualSIE/faces/javax.faces.resource/docs/Documento\\_operaciones\\_estadisticas\\_y\\_registros\\_administrativos.pdf](https://www.funcionpublica.gov.co/VisualSIE/faces/javax.faces.resource/docs/Documento_operaciones_estadisticas_y_registros_administrativos.pdf).
- Pizarro Hofer, R. (2001). La vulnerabilidad social y sus desafíos: una mirada desde América Latina. Cepal.
- Secretaría Distrital de Planeación – Analytica MSE SAS BIC (2023). Documento soporte del inventario de los programas distritales que requieren estratificación social. Mimeo.
- Secretaría Distrital de Planeación – Analytica MSE SAS BIC (2024). Diseño conceptual y metodológico para la construcción del Registro Universal de Ingresos. Mimeo.

## APÉNDICE 1. CONSTRUCCIÓN DEL ÍNDICE CON RUTINAS

### Presentación

El presente documento constituye un complemento a lo expuesto en el documento del producto 3 del proyecto, sobre el diseño del índice de ingreso del Registro Único de Ingresos (RUI) y la propuesta para su clasificación: *Diseño conceptual y metodológico para la construcción del registro universal de ingresos* (SDP, Analytica SAS, 2024).

La construcción de un índice de ingresos encuentra justificación en consideraciones de orden pragmático. Se busca a través de este instrumento tener una herramienta práctica que facilite el uso de la información, tanto para el análisis como para aplicar programas focalizados con base en información de ingresos. También es un recurso para tener mayor transparencia en el ejercicio de los programas sociales haciendo más comprensible la forma como operan y la norma de equidad que aplican ante la ciudadanía en general.

La segmentación o clasificación del índice de ingresos en grupos busca establecer segmentos poblacionales con características similares de ingresos que se tengan en cuenta en los procesos de focalización para garantiza en mejor forma que la focalización asegure, como criterio de equidad, que se dé tratamiento igual a quienes comparten características similares

### Construcción del Índice

El índice se construye en el ejercicio realizado sobre el ingreso al que se da prioridad, el ingreso corriente secundario per cápita, por unidad de gasto<sup>32</sup>, obtenido en la encuesta EMPB 2021, que es definido como:

$$Ing.CS_{it} = Ing.CP_{it} + T_{it}$$

Siendo:

$Ing.CS_{it}$ : Ingreso Corriente Secundario per cápita de la unidad i en el tiempo t

$Ing.CP_{it}$ : Ingreso Corriente Primario per cápita de la unidad i en el tiempo t

$T_{it}$ : Transferencias de la unidad i en el tiempo t

Para la construcción del índice de ingresos los valores del ingreso corriente secundario se transformaron a una escala que osciló entre 0 a 1, siguiendo los procedimientos aplicados en la construcción del Índice de Desarrollo Humano (IDH) en los Informes mundiales de

---

<sup>32</sup> Variable  $N\_ingpcugarr$  en los archivos digitales de la encuesta.

Desarrollo Humano realizados por el PNUD desde 1990 (PNUD, 1990 y 1992)<sup>33</sup>. Se escogió esta técnica debido a su sencillez y porque su aplicación cumple con los siguientes requisitos o condiciones<sup>34</sup>:

1. Que exprese adecuadamente la distribución de la variable ingreso que representa, obtenida de una fuente fiable.
2. Que permita una segmentación adecuada en estratos o grupos diferenciados por sus niveles de ingreso y sus condiciones de vida.
3. Que pueda ser actualizado periódicamente con economía de recursos.

En primer lugar, se hizo una transformación de la variable relativa a esos ingresos que simplifique su expresión numérica. Luego se asignó la unidad a todos aquellos registros donde el ingreso corriente secundario reportara valor cero (0); en seguida se calculó el logaritmo neperiano de la variable en cuestión (ya con el cambio de 0 a 1 para quienes carezcan de ingreso) y a continuación se obtuvieron el mínimo y máximo de la misma. Por último, se calculó el valor del índice para cada unidad de gasto a través de la siguiente ecuación:

$$IRUI_{it} = \frac{\ln(IngCS_{it}) - \ln(IngCS_{itmin})}{\ln(IngCS_{itmax}) - \ln(IngCS_{itmin})}$$

Donde:

$\ln(IngCS_{it})$  representa el logaritmo natural del ingreso secundario per-cápita de la unidad de gasto  $i$  en el tiempo  $t$ .

$\ln(IngCS_{itmax})$ : representa el valor máximo del logaritmo natural del ingreso corriente secundario del conjunto de unidades de gasto consideradas.

$\ln(IngCS_{itmin})$ : representa el valor mínimo del logaritmo del ingreso corriente secundario del conjunto de unidades de gasto consideradas.

El índice así construido expresa, en términos de proporción, la ubicación de cada individuo dentro del rango de la variable. Así, un valor de 0.1 significa que ese puntaje corresponde al 10% del valor máximo del ingreso expresado en logaritmos. Se usa el valor en logarítmicos para reducir la distancia entre valores superiores e inferiores y expresar el rendimiento decreciente del ingreso, a medida que aumenta. Ya que el logaritmo neperiano de cero es indefinido se le asigna un valor de 1 a quién no tiene ingresos ya que el logaritmo de 1 es igual a cero.

La rutina utilizada para este paso con el software IBM-SPSS-V24 ® fue la siguiente:

<sup>33</sup> Esta técnica fue propuesta y empezó a usarse con el Índice de Calidad de Vida Física (ICVF) elaborado por Morris, Morris D. (1979).

<sup>34</sup> Se hicieron ejercicios complementarios para aplicar otras técnicas para la construcción del índice como la incorporadas en algunas experiencias que hacen una *normalización* de las variables (véase, por ejemplo, Ibáñez, Castillo y Mojica, 2018) o la de utilizar medidas como la ubicación en cuantiles (percentiles, cuartiles, etc.) para clasificar los ingresos. Estas opciones se desecharon porque no expresaban adecuadamente la magnitud de las desigualdades de los ingresos.

```
WEIGHT BY FEX_C.  
RECODE N_ingpcugarr (0=1) (ELSE=Copy) INTO ingresopcu.  
EXECUTE.  
COMPUTE lningreso=LN(ingresopcu).  
EXECUTE.  
DESCRIPTIVES VARIABLES=lningreso  
/STATISTICS=MEAN STDDEV MIN MAX.  
COMPUTE IRui_Sec=(LN(ingresopcu)-MIN(LN(ingresopcu))/  
MAX(LN(ingresopcu)- MIN(LN(ingresopcu))).  
EXECUTE.
```

Luego se aplicaron cinco modelos diferentes de clasificación o segmentación en grupos del Índice de Ingresos corrientes secundarios obtenido en la encuesta EMPB 2021. Se buscó explorar opciones de clasificación en la perspectiva de establecer las que presentaran las mayores ventajas desde el punto de vista de su capacidad para representar grupos diferenciados según sus niveles de ingresos y sus condiciones de vida, y para facilitar la aplicación práctica de los procesos de focalización.

En los modelos se aplicaron distintas técnicas de clasificación o agrupamiento que son explicadas a continuación.

### Modelo 1.

En virtud de que a nivel personas, aproximadamente el 4.8% de la población se encuentra en unidades de gasto que tienen ingresos corrientes secundarios iguales a cero (EMP-2021) lo cual constituye una situación especial a la que importa dar un reconocimiento para la orientación de las políticas públicas<sup>35</sup>, y de que para fines prácticos convenía, también, identificar a través del índice a quienes se encuentran en situación de pobreza crítica (con ingresos per cápita inferiores a la línea de indigencia ) o no crítica (con ingresos per cápita inferiores a la línea de pobreza pero superiores a la línea de indigencia) se propuso el siguiente esquema para realizar la clasificación o segmentación de las unidades de gasto, a través del índice de ingresos:

Grupo 0: unidades cuyos ingresos per cápita sean cero (0).

Grupo 1: unidades por debajo de la línea de pobreza extrema e ingreso superior a cero.

Grupo 2: unidades que no sean pobres extremos pero que sí estén bajo la línea de la pobreza monetaria.

Para diferenciar los grupos del 3 al 6, se propuso utilizar una clasificación basada en algoritmo estadístico k-medias, sin especificar el centroide inicial del conglomerado. El algoritmo en cuestión procede de la siguiente manera:

---

<sup>35</sup> Véase, además, lo expuesto en relación con este tema y la medida de capacidad de pago, en los numerales 5.2.8 y 5.2.9. del documento del producto 3 (*Diseño conceptual y metodológico para la construcción del registro universal de ingresos*, SDP, Analytica SAS, 2024).



1. Se realizó una partición inicial en cuatro clústeres, con aquella población en Bogotá que no pertenezcan a los grupos 0, 1 y 2 antes definidos, con la finalidad de crear el centroide de cada uno de ellos. Esta fase inicial se realizó asignando inicialmente los casos más distantes entre sí a cada uno de los grupos.
2. Aquel individuo que estuviera más cerca (menor distancia) al centro de gravedad de un clúster particular sería reasignado a ese grupo. Sin embargo, hay que tener en cuenta que, en este método, un individuo asignado a un conglomerado en una determinada iteración puede ser reasignado a otro caso posteriormente.
3. Cada vez que se incorporaba un nuevo individuo al análisis, se ajustaron los nuevos centroides de los conglomerados.
4. Este procedimiento se repitió hasta reasignar toda la población en Bogotá a alguno de los cuatro grupos definidos, y que su distancia dentro de los clústeres fuera la mínima y entre ellos la máxima.

Una vez obtenido los cuatro grupos a través de la metodología de K-means, se tomaron las estadísticas descriptivas del índice propuesto de manera tal que permitieran asignar el orden del grupo correspondiente a cada conglomerado. Para esto se ordenaron los clústeres de la siguiente manera:

Sea  $K_{(j)}$ : el clúster ordenado según el valor máximo del índice de ingreso entre todos los grupos. Es decir:

- $K_{(1)}$  estarán aquella población en Bogotá que tuvieran los menores valores del índice, y son los que formarán el grupo 3.
- $K_{(2)}$  estará conformado por la población en Bogotá que tuvieran el índice de ingreso inmediatamente superior a  $K_{(1)}$ , y son los que formarán el grupo 4.
- $K_{(3)}$  estará conformado por la población en Bogotá que tuvieran el índice de ingreso inmediatamente superior a  $K_{(2)}$ , y son los que formarán el grupo 5.
- $K_{(4)}$  estará conformado por la población en Bogotá que tuvieran el índice de ingreso inmediatamente superior a  $K_{(3)}$ , y son los que formarán el grupo 6.
- $K_{(6)}$  estará conformado por la población en Bogotá que tuvieran el mayor índice de ingreso.

Por lo que adicionando los tres primeros grupos con las características antes mencionadas y los obtenidos por los conglomerados se tiene la siguiente clasificación (Tabla A1-1):

*Tabla A1-1. Grupos de clasificación según el Modelo propuesto 1*

Grupo	Descripción
G1_0	Índice de Ingresos 0
G1_1	Índice para aquellos ingresos por debajo de la línea de indigencia
G1_2	Índice para aquellos ingresos por debajo de la línea de la pobreza monetaria que no sean pobres extremos
G1_3	Índices de ingresos menores o iguales al del clúster $K_{(1)}$

Grupo	Descripción
G1_4	Índices de ingresos mayores al clúster $K_{(1)}$ y menores o iguales al clúster $K_{(2)}$
G1_5	Índices de ingresos mayores al clúster $K_{(2)}$ y menores o iguales al clúster $K_{(3)}$
G1_6	Índices de ingresos mayores al clúster $K_{(3)}$ y menores o iguales al clúster $K_{(4)}$

*Fuente: elaboración propia*

Se utilizó el algoritmo k-medias ya que este trata de minimizar el criterio clásico de distancia de cada una de las observaciones de la muestra a la media de su grupo al cuadrado. Dicho algoritmo presenta una serie de extensiones y generalizaciones entre las que destacan el caso borroso o fuzzy, el de máxima verosimilitud, y aquellos criterios basados en convexidad, entre otros.

Este análisis comienza a tomar su importancia desde la publicación de Sokal, (1963). Jardine (1971) profundiza en el desarrollo matemático del problema del agrupamiento, luego Anderberg (1970) se enfoca más hacia las aplicaciones, Sodeur (2013) utiliza métodos empíricos necesarios para realizar actividades de clasificación. A partir de estos estudios y otros más, este tipo de análisis de agrupamiento se volvieron más conocidos por la comunidad científica.

Tal como lo señala Valera (2017), uno de los principales enfoques de las técnicas de agrupamiento se basa en el criterio de la suma de los cuadrados de la varianza, y en el algoritmo que hoy en día se denomina k-medias. Con esta técnica se logró, a través de un proceso iterativo, que se conformarán grupos lo más homogéneos a su interior y lo más diferentes con los otros grupos.

La rutina utilizada para este paso utilizando el software IBM-SPSS-V24 ® fue la siguiente:

\*Modelo 1

Para crear los tres primeros grupos se realizó la siguiente rutina.

```
IF (N_ingpcugarr=0) Modelo_1=0.
EXECUTE.
IF (N_ingpcugarr>0 & N_pobre_extremo=1) Modelo_1=1.
EXECUTE.
IF (N_pobre_monetario=1 & N_pobre_extremo=0) Modelo_1=2.
EXECUTE.
```

Luego se creó una variable ficticia (Ficticia\_1) para seleccionar aquellos individuos que entraban a ser clasificados por el algoritmo de clúster:

```
RECODE Modelo_1 (0=0) (1=0) (2=0) (ELSE=1) INTO Ficticia_1.
VARIABLE LABELS Ficticia_1 'Variable ficticia para la creación de los clústeres
Modelo 1'.
```

EXECUTE.

Una vez creada esta variable se seleccionaron aquellos que tuvieran valor en la variable Ficticia\_1=1 a través de la siguiente rutina:

```
USE ALL.  
COMPUTE filter_$=(Ficticia_1 = 1).  
VARIABLE LABELS filter_$ 'Ficticia_1 = 1 (FILTER)'.  
VALUE LABELS filter_$ 0 'Not Selected' 1 'Selected'.  
FORMATS filter_$ (f1.0).  
FILTER BY filter_$.  
EXECUTE.
```

Ahora bien, con aquellos que tuvieron valor en la variable Ficticia\_1=1, se realizó el algoritmo de clúster con la siguiente rutina:

```
QUICK CLUSTER IRui_Sec  
/MISSING=LISTWISE  
/CRITERIA=CLUSTER(4) MXITER(10) CONVERGE(0)  
/METHOD=KMEANS(NOUPDATE)  
/SAVE CLUSTER  
/PRINT INITIAL.  
/PRINT INITIAL ANOVA.
```

Con los valores obtenidos en los cuatro clústeres, se les calcularon las estadísticas descriptivas del IRio\_Sec para ordenar de manera ascendente los grupos, a través de la siguiente rutina:

```
SORT CASES BY QCL_1.  
SPLIT FILE LAYERED BY QCL_1.  
DESCRIPTIVES VARIABLES=IRui_Sec  
/STATISTICS=MEAN STDDEV MIN MAX.
```

Una vez obtenidos los valores a asignar a cada uno de los grupos, se integraron a la variable correspondiente al Modelo\_1 para así obtener la variable de clasificación. Para esto se utilizó la siguiente rutina:

```
IF (QCL_1=2) Modelo_1=3.  
EXECUTE.  
IF (QCL_1=1) Modelo_1=4.  
EXECUTE.  
IF (QCL_1=3) Modelo_1=5.  
EXECUTE.  
IF (QCL_1=4) Modelo_1=6.  
EXECUTE.  
FILTER OFF.  
USE ALL.
```

```
EXECUTE.  
SORT CASES BY Modelo_1.  
SPLIT FILE LAYERED BY Modelo_1.  
SPLIT FILE OFF.
```

Como prueba de la bondad del ajuste de las clasificaciones efectuadas se recurrió a la Prueba no paramétrica de Kruskal-Wallis la cual se utiliza para determinar si más de dos muestras provienen de la misma distribución y para probar la igualdad de más de dos poblaciones. Así las cosas, cuando la distribución de los datos no es normal, se asume bajo la hipótesis nula que los datos provienen de la misma población. De igual manera, evalúa si las medianas de la población sobre una variable dependiente, en este caso el RUI, son las mismas para todos los niveles de clasificación obtenidos.

La hipótesis nula planteada es que las muestras provienen de la misma población continua. A su vez, la hipótesis alternativa es que al menos una de las poblaciones tiende a producir comparativamente valores más grandes que las otras (Green, 2008)

Si  $R_i$  es la suma de las filas de los datos en la muestra  $i$ -ésima de tamaño  $n_i$  y  $n$  es la suma total de las muestras  $k$ , la estadística de prueba normalizado es:

$$H = \frac{12}{n(n+1)} \sum_{i=1}^k \frac{R_i^2}{n_i} - 3(n+1)$$

Donde  $k$  es el número de grupos.

Bajo la hipótesis nula de que las muestras provienen de la misma población,  $H$  tiene un aproximado de distribución chi-cuadrado con  $(k-1)$  grados de libertad.

La rutina que permitió realizar esta prueba para el grupo correspondiente es:

```
*Nonparametric Tests: Independent Samples.  
NPTESTS  
/INDEPENDENT TEST (IRui_Sec) GROUP (Modelo_1)  
/MISSING SCOPE=ANALYSIS USERMISSING=EXCLUDE  
/CRITERIA ALPHA=0.05 CILEVEL=95.
```

## Modelo 2

Para este modelo se propusieron seis grupos, usando la metodología que se expone a continuación. Los dos primeros grupos estarían formados por aquellas unidades ubicadas por debajo de la pobreza extrema, y por aquellos que no sean indigentes, pero que si estén clasificados como pobres monetarios respectivamente.

Para diferenciar los grupos del 3 al 6, se propuso utilizar una clasificación basada en algoritmo estadístico k-means, sin especificar el centroide del conglomerado. El algoritmo en cuestión procedió de la siguiente manera:

- Se realizó una primera partición en cuatro clústeres, con aquella población en Bogotá que no pertenecían a los grupos 1 y 2 antes definidos, con la finalidad de crear el centroide de cada uno de ellos. Esta fase inicial se realizó asignando inicialmente los casos más distantes entre sí a cada uno de los grupos.
- Aquella unidad que estuviera más cerca (menor distancia) al centro de gravedad de un clúster particular sería reasignada a ese grupo. Sin embargo, hay que tener en cuenta que, en este método, una unidad asignada a un conglomerado en una determinada iteración puede ser reasignado a otro caso posteriormente.
- Cada vez que se incorporó una nueva unidad, se ajustaron los centroides de cada uno de los conglomerados.
- Este procedimiento se repitió hasta reasignar todas las unidades a alguno de los cuatro grupos definidos, y que su distancia dentro de los clústeres fuera la mínima y entre ellos la máxima.

Una vez obtenido los cuatro grupos a través de la metodología de K-means, se tomaron las estadísticas descriptivas del índice propuesto de manera tal que permitieran asignar el grupo correspondiente a cada conglomerado. Para esto se ordenaron los clústeres de la siguiente manera:

Sea  $K_{(j)}$ : el clúster ordenado según el valor máximo del índice de ingreso entre todos los grupos. Es decir:

- $K_{(1)}$  estarán aquellas unidades que tengan los menores valores del índice, y se les asignará el Grupo 3.
- $K_{(2)}$  estará conformado por las unidades que tengan el índice de ingreso inmediatamente superior a  $K_{(1)}$ , y se les asignará el Grupo 4.
- $K_{(3)}$  lo conformarán las unidades que tengan el índice de ingreso inmediatamente superior a  $K_{(2)}$ , y se les asignará el Grupo 5.
- $K_{(4)}$  estará conformado por las unidades que tengan el mayor índice de ingreso y se les asignará el Grupo 6.

Por lo que adicionando los dos primeros grupos con las características antes mencionadas y los obtenidos por los conglomerados se obtuvo la siguiente clasificación (Tabla A1-2):

*Tabla A1-2. Grupos de clasificación según el Modelo propuesto 2*

Grupos Modelo 2	Descripción
G2_1	Índice para aquellos ingresos por debajo de la línea de indigencia
G2_2	Índice para aquellos ingresos por debajo de la línea de la pobreza monetaria que no sean pobres extremos
G2_3	Índices de ingresos menores o iguales al del clúster $K_{(1)}$
G2_4	Índices de ingresos mayores al clúster $K_{(1)}$ y menores o iguales al clúster $K_{(2)}$
G2_5	Índices de ingresos mayores al clúster $K_{(2)}$ y menores o iguales al clúster $K_{(3)}$
G2_6	Índices de ingresos mayores al clúster $K_{(3)}$ y menores o iguales al clúster $K_{(4)}$

*Fuente: elaboración propia*

La rutina utilizada para este paso utilizando el software IBM-SPSS-V24 ® fue la siguiente:

\*Modelo 2

Para crear los tres primeros grupos se realizó la siguiente rutina:

```
IF (N_ingpcugarr>0 & N_pobre_extremo=1) Modelo_2=1.  
EXECUTE.  
IF (N_pobre_monetario=1 & N_pobre_extremo=0) Modelo_2=2.  
EXECUTE.
```

Luego se creó una variable ficticia (Ficticia\_2) para seleccionar aquellos individuos que entraban a ser clasificados por el algoritmo de clúster:

```
RECODE Modelo_2 (0=0) (1=0) (2=0) (ELSE=1) INTO Ficticia_2.  
VARIABLE LABELS Ficticia_2 'Variable ficticia para la creación de los clústeres  
Modelo 1'.  
EXECUTE.
```

Una vez creada esta variable se seleccionaron aquellos que tuvieran valor en la variable Ficticia\_2=1 a través de la siguiente rutina:

```
USE ALL.  
COMPUTE filter_$=(Ficticia_2 = 1).  
VARIABLE LABELS filter_$ 'Ficticia_2 = 1 (FILTER)'.  
VALUE LABELS filter_$ 0 'Not Selected' 1 'Selected'.  
FORMATS filter_$ (f1.0).  
FILTER BY filter_$.  
EXECUTE.
```

Ahora bien, con aquellos que tuvieron valor en la variable Ficticia\_2=1, se realizó el algoritmo de clúster con la siguiente rutina:

```
QUICK CLUSTER IRui_Sec  
/MISSING=LISTWISE  
/CRITERIA=CLUSTER(4) MXITER(10) CONVERGE(0)  
/METHOD=KMEANS(NOUPDATE)  
/SAVE CLUSTER  
/PRINT INITIAL.  
/PRINT INITIAL ANOVA.
```

Con los valores obtenidos en los cuatro clústeres, se les calcularon las estadísticas descriptivas del IRio\_Sec para ordenar de manera ascendente los grupos, a través de la siguiente rutina:

```
SORT CASES BY QCL_2.  
SPLIT FILE LAYERED BY QCL_2.  
DESCRIPTIVES VARIABLES=IRui_Sec
```



/STATISTICS=MEAN STDDEV MIN MAX.

Una vez obtenidos los valores a asignar a cada uno de los grupos, se integraron a la variable correspondiente al Modelo\_2 para así obtener la variable de clasificación. Para esto se utilizó la siguiente rutina:

```
IF (QCL_2=3) Modelo_2=3.  
EXECUTE.  
IF (QCL_2=4) Modelo_2=4.  
EXECUTE.  
IF (QCL_2=2) Modelo_2=5.  
EXECUTE.  
IF (QCL_2=1) Modelo_2=6.  
EXECUTE.  
FILTER OFF.  
USE ALL.  
EXECUTE.  
SORT CASES BY Modelo_2.  
SPLIT FILE LAYERED BY Modelo_2.  
SPLIT FILE OFF.
```

La rutina que permitió realizar esta prueba de Kruskal Wallis para el grupo correspondiente fue:

```
NPTESTS  
/INDEPENDENT TEST (IRui_Sec) GROUP (Modelo_2)  
/MISSING SCOPE=ANALYSIS USERMISSING=EXCLUDE  
/CRITERIA ALPHA=0.05 CILEVEL=95.
```

### Modelo 3

Se propuso, al igual que en el modelo 2, una clasificación en seis grupos, con la diferencia que todos son seleccionados según el algoritmo k-means, sin detallar el centroide de cada uno de ellos.

- Se realizó una primera partición en seis grupos, con la finalidad de crear el centroide de cada uno de ellos. Esta fase inicial se realizó asignando primero los casos más distantes entre sí a cada uno de los grupos.
- Aquella unidad que estuviera a menor distancia del centro de gravedad de un grupo particular sería reasignada a ese clúster. Se debió tener en cuenta que una unidad atribuida a un conglomerado en una iteración determinada podía ser reasignado a otro caso posteriormente.
- A medida que se incorpora una nueva unidad, se ajustaban los centroides de los grupos.

- Este procedimiento se repitió hasta reasignar todas las unidades a alguno de los seis grupos definidos, y que su distancia dentro de los ellos fuera la mínima y entre ellos la máxima.

Obtenidos los seis grupos a través de ésta, se calcularon las estadísticas descriptivas del índice propuesto para asignar el grupo correspondiente a cada conglomerado. Para esto se ordenaron los clústeres de la siguiente manera:

Sea  $K_{(j)}$ : el clúster ordenado según el valor máximo del índice de ingreso entre todos los grupos. Es decir:

- $K_{(1)}$  estarán aquellas unidades que tengan los menores valores del índice, y se les asignará el Grupo 1.
- $K_{(2)}$  estará conformado por las unidades que tengan el índice de ingreso inmediatamente superior a  $K_{(1)}$ , y se les asignará el Grupo 1.
- $K_{(3)}$  lo conformarán las unidades que tengan el índice de ingreso inmediatamente superior a  $K_{(2)}$ , y se les asignará el Grupo 3.
- $K_{(4)}$  estará conformado por las unidades que tengan el índice de ingreso inmediatamente superior a  $K_{(3)}$ , y se les asignará el Grupo 4.
- $K_{(5)}$  estará conformado por las unidades que tengan el índice de ingreso inmediatamente superior a  $K_{(4)}$ , y se les asignará el Grupo 5.
- $K_{(6)}$  estará formado por las unidades que tengan el mayor índice de ingreso y se les asignará el Grupo 6.

Por lo tanto, los grupos de clasificación de ingresos según el Modelo 3 son (Tabla A1-3):

*Tabla A1-3. Grupos de clasificación según el Modelo propuesto 3*

Grupos Modelo 3	Descripción
G3_1	Índices de ingresos menores o iguales al del clúster $K_{(1)}$
G3_2	Índices de ingresos mayores al clúster $K_{(1)}$ y menores o iguales al clúster $K_{(2)}$
G3_3	Índices de ingresos mayores al clúster $K_{(2)}$ y menores o iguales al clúster $K_{(3)}$
G3_4	Índices de ingresos mayores al clúster $K_{(3)}$ y menores o iguales al clúster $K_{(4)}$
G3_5	Índices de ingresos mayores al clúster $K_{(4)}$ y menores o iguales al clúster $K_{(5)}$
G3_6	Índices de ingresos mayores al clúster $K_{(5)}$

*Fuente: elaboración propia*

La rutina utilizada para este paso utilizando el software IBM-SPSS-V24 ® fue la siguiente:

```
QUICK CLUSTER Irui_Sec
/MISSING=LISTWISE
/CRITERIA=CLUSTER(6) MXITER(10) CONVERGE(0)
/METHOD=KMEANS(NOUPDATE)
/SAVE CLUSTER
/PRINT INITIAL ANOVA.
```

```
SORT CASES BY QCL_3.
SPLIT FILE LAYERED BY QCL_3.
```

```
DESCRIPTIVES VARIABLES=IRui_Sec N_ingpcugarr
/STATISTICS=MEAN STDDEV MIN MAX.
```

```
RECODE QCL_3 (1=1) (2=5) (3=4) (4=6) (5=2) (6=3) INTO Modelo_3.
VARIABLE LABELS Modelo_3 'Grupos Modelo 3'.
EXECUTE.
```

La rutina que permitió realizar la prueba de Kruskal Wallis para el grupo correspondiente fue:

```
NPTESTS
/INDEPENDENT TEST (IRui_Sec) GROUP (Modelo_3)
/MISSING SCOPE=ANALYSIS USERMISSING=EXCLUDE
/CRITERIA ALPHA=0.05 CILEVEL=95.
```

#### Modelo 4

Para este modelo se propuso dividir el conjunto de datos en seis grupos aproximadamente iguales en cuanto al número de entidades que irían en cada uno de ellos. Para esto se tomaron los valores de los índices correspondientes a los percentiles: 16.67, 33.33, 50, 66.67 y 83.33. Así las cosas, la clasificación correspondiente sería (Tabla A1-4):

*Tabla A1-4. Grupos de clasificación según el Modelo propuesto 4*

Grupos Modelo 4	Descripción
G4_1	Índices de ingresos menores al valor del percentil 16.67
G4_2	Índices de ingresos entre los valores de los percentiles 16.67 y 33.33
G4_3	Índices de ingresos entre los valores de los percentiles 33.33 y 50
G4_4	Índices de ingresos entre los valores de los percentiles 50 y 66.67
G4_5	Índices de ingresos entre los valores de los percentiles 66.67 y 83.33
G4_6	Índices de ingresos mayores al valor del percentil 83.33

*Fuente: elaboración propia*

La rutina que permitió realizar este agrupamiento fue:

```
RECODE Irui_Sec (Lowest thru 0.66751966685523=1) (0.66751966685523 thru
0.69908715691442=2)
(0.69908715691442 thru 0.723020295379983=3) (0.723020295379983 thru
0.747562851327629=4)
(0.747562851327629 thru 0.782723827169989=5) (0.782723827169989 thru
Highest=6) INTO Modelo_4.
EXECUTE.
```

La rutina que permitió realizar la prueba de Kruskal Wallis para el grupo correspondiente fue:

## NPTESTS

```
/INDEPENDENT TEST (IRui_Sec) GROUP (Modelo_4)
/MISSING SCOPE=ANALYSIS USERMISSING=EXCLUDE
/CRITERIA ALPHA=0.05 CILEVEL=95.
```

## Modelo 5

Siguiendo una conceptualización similar a la propuesta en el Modelo 4, para este modelo se proponen dividir el conjunto de datos según los cuartiles correspondientes. En este orden de ideas, cada grupo tendrá aproximadamente el 25% de las entidades bajo estudio. Para esto se toman los valores de los índices correspondientes a los cuartiles 1, 2 y 3. En este orden de ideas, la clasificación correspondiente sería (Tabla A1-5):

*Tabla A1-5. Grupos de clasificación según el Modelo propuesto 5*

Grupos Modelo 5	Descripción
G5_1	Índices de ingresos menores al valor del cuartil 1
G5_2	Índices de ingresos entre los valores de los cuartiles 1 y 2
G5_3	Índices de ingresos entre los valores de los cuartiles 2 y 3
G5_4	Índices de ingresos mayores al valor del cuartil 3

*Fuente: elaboración propia*

Las rutinas que permitieron realizar esta clasificación, así como la prueba de Kruskal Wallis fueron:

```
RECODE Irui_Sec (Lowest thru 0.684946951794784=1) (0.684946951794784 thru
0.723020295379983=2)
(0.723020295379983 thru 0.763747350487242=3) (0.763747350487242 thru
Highest=4) INTO Modelo_5.
EXECUTE.
```

*\*Nonparametric Tests: Independent Samples.*

## NPTESTS

```
/INDEPENDENT TEST (IRui_Sec) GROUP (Modelo_5)
/MISSING SCOPE=ANALYSIS USERMISSING=EXCLUDE
/CRITERIA ALPHA=0.05 CILEVEL=95.
```

Una vez obtenido los cinco modelos, y en virtud que el mayor valor de la chi-cuadrado con 5 g.l. lo obtuvo el Modelo 2, se consideró, bajo este punto de vista, como el mejor para establecer los grupos de clasificación desde esta perspectiva. No obstante, como se mencionó en otras partes del Producto 3, el modelo 1 tiene ventajas desde otros puntos de vista.

Ahora bien, para comprobar la capacidad de los grupos de los modelos para captar segmentos diferenciados desde el punto de vista de sus ingresos, se aplicó adicionalmente la técnica de regresión logística multinomial (RLM) a los cinco modelos propuestos para la clasificación del índice de ingresos corrientes secundarios obtenidos en la encuesta EMPB 2021. Con el fin de evaluar la capacidad de los modelos para clasificar adecuadamente a los hogares en los grupos predefinidos, se especificaron los modelos de regresión fijando como variable respuesta la clasificación en cada grupo y como covariable el índice de ingresos. Para esto se realizaron las siguientes rutinas para cada uno de los modelos antes descritos:

#### Modelo 1

```
NOMREG Modelo_1 (BASE=LAST ORDER=ASCENDING) BY Irui_Sec
/CRITERIA CIN(95) DELTA(0) MXITER(100) MXSTEP(5) CHKSEP(20)
LCONVERGE(0) PCONVERGE(0.000001)
SINGULAR(0.00000001)
/MODEL
/STEPWISE=PIN(.05) POUT(0.1) MINEFFECT(0) RULE(SINGLE)
ENTRYMETHOD(LR) REMOVALMETHOD(LR)
/INTERCEPT=INCLUDE
/PRINT=CLASSTABLE FIT PARAMETER SUMMARY LRT CPS STEP MFI.
```

#### Modelo 2

```
NOMREG Modelo_2 (BASE=LAST ORDER=ASCENDING) BY Irui_Sec
/CRITERIA CIN(95) DELTA(0) MXITER(100) MXSTEP(5) CHKSEP(20)
LCONVERGE(0) PCONVERGE(0.000001)
SINGULAR(0.00000001)
/MODEL
/STEPWISE=PIN(.05) POUT(0.1) MINEFFECT(0) RULE(SINGLE)
ENTRYMETHOD(LR) REMOVALMETHOD(LR)
/INTERCEPT=INCLUDE
/PRINT=CLASSTABLE FIT PARAMETER SUMMARY LRT CPS STEP MFI.
```

#### Modelo 3

```
NOMREG Modelo_3 (BASE=LAST ORDER=ASCENDING) BY Irui_Sec
/CRITERIA CIN(95) DELTA(0) MXITER(100) MXSTEP(5) CHKSEP(20)
LCONVERGE(0) PCONVERGE(0.000001)
SINGULAR(0.00000001)
/MODEL
/STEPWISE=PIN(.05) POUT(0.1) MINEFFECT(0) RULE(SINGLE)
ENTRYMETHOD(LR) REMOVALMETHOD(LR)
/INTERCEPT=INCLUDE
/PRINT=CLASSTABLE FIT PARAMETER SUMMARY LRT CPS STEP MFI.
```

#### Modelo 4

```
NOMREG Modelo_4 (BASE=LAST ORDER=ASCENDING) BY Irui_Sec
/CRITERIA CIN(95) DELTA(0) MXITER(100) MXSTEP(5) CHKSEP(20)
LCONVERGE(0) PCONVERGE(0.000001)
SINGULAR(0.00000001)
/MODEL
/STEPWISE=PIN(.05) POUT(0.1) MINEFFECT(0) RULE(SINGLE)
ENTRYMETHOD(LR) REMOVALMETHOD(LR)
/INTERCEPT=INCLUDE
/PRINT=CLASSTABLE FIT PARAMETER SUMMARY LRT CPS STEP MFI.
```

#### Modelo 5

```
NOMREG Modelo_5 (BASE=LAST ORDER=ASCENDING) BY Irui_Sec
/CRITERIA CIN(95) DELTA(0) MXITER(100) MXSTEP(5) CHKSEP(20)
LCONVERGE(0) PCONVERGE(0.000001)
SINGULAR(0.00000001)
/MODEL
/STEPWISE=PIN(.05) POUT(0.1) MINEFFECT(0) RULE(SINGLE)
ENTRYMETHOD(LR) REMOVALMETHOD(LR)
/INTERCEPT=INCLUDE
/PRINT=CLASSTABLE FIT PARAMETER SUMMARY LRT CPS STEP MFI.
```

Se presenta a continuación los resultados de la RLM aplicada a esos cinco modelos.

Se obtuvo que todos los ellos eran significativos, teniendo en cuenta los resultados de las pruebas de bondad de ajuste que, como se expresa en la tabla siguiente, presentan p-value con valores menores a 0.05 (Tabla A1-6). Esto significa que la variable utilizada como covariable explica adecuadamente la variable respuesta. Además, los pseudo-r cuadrado calculados para los modelos son mayores al 0.95, indicando que, en una proporción importante, la ubicación en los grupos está explicada por el índice de ingresos.

*Tabla A1-6. Ajuste de los modelos de regresión logística multinomial*

Información de ajuste de los modelos					
Modelos de Regresión Logística Multinomial		Criterios de ajuste de modelo	Pruebas de la razón de verosimilitud	p-value	
		Logaritmo de la verosimilitud -2	Chi-cuadrado	gl	Sig.
Modelo 1	Sólo intersección	27.268.829			
	Final	559.435	26.709.394	6	0,00
Modelo 2	Sólo intersección	25.951.974			
	Final	563.465	25.388.509	5,00	0,00



Información de ajuste de los modelos					
Modelos de Regresión Logística Multinomial		Criterios de ajuste de modelo	Pruebas de la razón de verosimilitud	p-value	
		Logaritmo de la verosimilitud -2	Chi-cuadrado	gl	Sig.
<b>Modelo 3</b>	Sólo intersección	24.099.969			
	Final	14.348	24.085.621	5,00	0,00
<b>Modelo 4</b>	Sólo intersección	28.153.685			
	Final	39.865	28.113.819	5	0,00
<b>Modelo 5</b>	Sólo intersección	21.784.732			
	Final	39.057	21.745.675	3,00	0,00

*Fuente: elaboración propia a partir de los resultados EMP 2021*

La RLM también permitió constatar que una alta proporción de las predicciones son correctas, como se muestra en las tablas A1-7 a A1-11. En ellas se observa que en la diagonal se ubican las predicciones correctas y las observaciones que quedan fuera de la diagonal son incorrectas. En todos los modelos los porcentajes totales de aciertos en las clasificaciones son superiores a 98.75%. Esto da base para concluir que las clasificaciones de los cinco modelos de clasificación son adecuadas, a niveles que pueden considerarse equivalentes de acuerdo con la prueba realizada. Aunque las diferencias en el porcentaje de aciertos son poco significativas entre esos modelos de clasificación, los que tienen valores un poco más altos, próximos al 100%, son los modelos 3, 4 y 5.

*Tabla A1-7. Casos pronosticados y observados según la regresión logística multinomial del modelo de clasificación 1.*

<b>Modelo 1</b>								
Observado	Pronosticado							Porcentaje correcto
	Grupo 0	Grupo 1	Grupo 2	Grupo 3	Grupo 4	Grupo 5	Grupo 6	
<b>Grupo 0</b>	379.112	0	0	0	0	0	0	100%
<b>Grupo 1</b>	0	566.653	43.399	0	0	0	0	92,89%
<b>Grupo 2</b>	0	10.549	1.752.538	18.062	0	0	0	98,39%
<b>Grupo 3</b>	0	0	20.065	2.191.274	0	0	0	99,09%
<b>Grupo 4</b>	0	0	0	4.057	1.579.113	291	0	99,73%
<b>Grupo 5</b>	0	0	0	0	0	959.224	0	100%
<b>Grupo 6</b>	0	0	0	0	0	21	333.096	99,99%
								<b>98,77%</b>

*Fuente: elaboración propia a partir de los resultados EMP 2021*

Tabla A1-8. Casos pronosticados y observados según la regresión logística multinomial del modelo de clasificación 2.

<b>Modelo 2</b>							
<b>Observado</b>	<b>Pronosticado</b>						<b>Porcentaje correcto</b>
	<b>Grupo 1</b>	<b>Grupo 2</b>	<b>Grupo 3</b>	<b>Grupo 4</b>	<b>Grupo 5</b>	<b>Grupo 6</b>	
<b>Grupo 1</b>	945.765	43.399	0	0	0	0	95,61%
<b>Grupo 2</b>	10.549	1.752.538	18.062	0	0	0	98,39%
<b>Grupo 3</b>	0	20.065	2.191.274	0	0	0	99,09%
<b>Grupo 4</b>	0	0	4.418	1.578.320	723	0	99,68%
<b>Grupo 5</b>	0	0	0	0	959.224	0	100%
<b>Grupo 6</b>	0	0	0	0	88	333.029	99,97%
							<b>98,76%</b>

Fuente: elaboración propia a partir de los resultados EMP 2021

Tabla A1-9. Casos pronosticados y observados según la regresión logística multinomial del modelo de clasificación 3.

<b>Modelo 3</b>							
<b>Observado</b>	<b>Pronosticado</b>						<b>Porcentaje correcto</b>
	<b>Grupo 1</b>	<b>Grupo 2</b>	<b>Grupo 3</b>	<b>Grupo 4</b>	<b>Grupo 5</b>	<b>Grupo 6</b>	
<b>Grupo 1</b>	379.112	0	0	0	0	0	100%
<b>Grupo 2</b>	0	282.577	22	0	0	0	99,99%
<b>Grupo 3</b>	0	0	2.049.436	1.670	0	0	99,92%
<b>Grupo 4</b>	0	0	0	2.724.200	0	0	100%
<b>Grupo 5</b>	0	0	0	451	1.703.339	0	99,97%
<b>Grupo 6</b>	0	0	0	0	0	716.648	100%
							<b>99,97%</b>

Fuente: elaboración propia a partir de los resultados EMP 2021

*Tabla A1-10. Casos pronosticados y observados según la regresión logística multinomial del modelo de clasificación 4.*

<b>Modelo 4</b>							
<b>Observado</b>	<b>Pronosticado</b>						<b>Porcentaje correcto</b>
	<b>Grupo 1</b>	<b>Grupo 2</b>	<b>Grupo 3</b>	<b>Grupo 4</b>	<b>Grupo 5</b>	<b>Grupo 6</b>	
<b>Grupo 1</b>	1.328.278	0	0	0	0	0	100%
<b>Grupo 2</b>	1.928	1.251.268	1.507	0	0	0	99,73%
<b>Grupo 3</b>	0	0	1.346.926	0	0	0	100%
<b>Grupo 4</b>	0	0	81	1.308.322	0	0	99,99%
<b>Grupo 5</b>	0	0	0	0	1.309.610	0	100%
<b>Grupo 6</b>	0	0	0	0	410	1.309.126	99,97%
							<b>99,95%</b>

*Fuente: elaboración propia a partir de los resultados EMP 2021*

*Tabla A1-11. Casos pronosticados y observados según la regresión logística multinomial del modelo de clasificación 1.*

<b>Modelo 5</b>					
<b>Observado</b>	<b>Pronosticado</b>				<b>Porcentaje correcto</b>
	<b>Grupo 1</b>	<b>Grupo 2</b>	<b>Grupo 3</b>	<b>Grupo 4</b>	
<b>Grupo 1</b>	1.964.795	0	0	0	100%
<b>Grupo 2</b>	531	1.964.580	0	0	99,97%
<b>Grupo 3</b>	0	641	1.990.416	0	99,97%
<b>Grupo 4</b>	0	0	3.616	1.932.874	99,81%
					<b>99,94%</b>

*Fuente: elaboración propia a partir de los resultados EMP 2021*

Para finalizar se presenta el balance de las pruebas aplicadas para evaluar los modelos de clasificación propuestos:

Los resultados obtenidos y las pruebas aplicadas en los cinco modelos sobre la capacidad de los grupos escogidos para diferenciarse adecuadamente en términos de niveles de ingresos permiten concluir que las segmentaciones en cada uno de ellos son consistentes y permiten llegar a delimitaciones apropiadas de las unidades de gasto según la variable principal de interés: el ingreso secundario per cápita. Sobre esta base se considera que no se justifica, continuar con ejercicios exploratorio-adicionales utilizando otras metodologías con la Encuesta Multipropósito 2021. Lo obtenido cumple con las cualidades esperadas y

permite, además, establecer que el agrupamiento del modelo 2 presenta ventajas en términos de sus propiedades estadísticas discriminatorias frente a los otros modelos.

Desde el punto de vista metodológico la técnica de clasificación estadística *k-means* (modelos 1, 2 y 3) ofrece ventajas frente a la segmentación de cuantiles (modelos 5 y 6). Aunque con todos los métodos se logra diferenciar grupos homogéneos bajo el criterio de los ingresos, desde la perspectiva de la focalización es más importante conseguir una mejor segmentación bajo este concepto, lograda más adecuadamente con esta técnica estadística, que obtener grupos con tamaño similar (quintiles, rangos de percentiles), con lo cual se alcanza menos homogeneidad entre los grupos. Como opción para efectuar la focalización, la primera metodología garantiza en mejor forma que la focalización asegure, como criterio de equidad, que se dé tratamiento igual a los que comparten características similares.

Las agrupaciones de los modelos 1 y 2, introducen, de otra parte, que se considere la ubicación de la población en las categorías de pobreza crítica y moderada. Así se garantizan condiciones para llegar a resultados óptimos en términos de minimizar los errores de inclusión y exclusión, cumpliendo con un criterio de evaluación decisivo en términos de los atributos requeridos para una clasificación de los ingresos del RUI. Es este otro argumento que contribuye a ratificar la validez de las ventajas de las clasificaciones hecha en estos dos modelos.

Adicionalmente, la comprobación, que será expuesta en el documento del producto 4, de la capacidad de los grupos definidos de distinguirse no solo por sus niveles de ingreso, sino también de acuerdo con características demográficas y socio-económicas particulares, es también un criterio de validación adicional a favor de las clasificaciones propuestas.

### **K medias**

El algoritmo k medias trata de minimizar el criterio clásico de distancia de cada una de las observaciones de la muestra a la media de su grupo al cuadrado.

### **Bondad de ajuste**

Como prueba de la Bondad del ajuste se utilizó la Prueba no paramétrica de Kruskal-Wallis, la cual se utiliza para determinar si más de dos muestras que provienen de la misma distribución son iguales.

La hipótesis nula planteada es que las muestras que provienen de la misma población continua son iguales. La hipótesis alternativa es que al menos una de las poblaciones tiende a producir comparativamente valores más grandes que las otras.

### **Una prueba adicional sobre la adecuación de la clasificación propuesta con la técnica de regresión logística**

Para comprobar la capacidad de los grupos de los modelos para captar segmentos diferenciados desde el punto de vista de sus ingresos, se aplicó adicionalmente la técnica de regresión logística multinomial (RLM) a los cinco modelos propuestos para la clasificación del índice de ingresos corrientes secundarios obtenidos en la encuesta EMB 2021 (Tabla A1-12).

Tabla A1-12. Modelo de clasificación RUI – Comparaciones

Modelo	Nº de grupos de clasificación	Prueba estadística										Pobreza extrema		Pobreza Monetaria	
		Kruskal-Wallis						Modelo de Regresión Multinomial				Error Inclusión	Error Exclusión	Error Inclusión	Error Exclusión
		Estadístico H	N	Valor chi-cuadrado	Grados de libertad	p-valor	Tamaño del efecto	Log de la verosimilitud -2	Valor chi-cuadrado	p-valor	Porcentaje de clasificación				
1	7 grupos	7.472.985,99	7.858.011	7.499.509,02	6	0,000	0,00035	559.435	26.709.394	0,000	98,77%	25,93%	0	0	0
2	6 grupos	7.513.617,29	7.858.011	7.488.262,31	5	0,000	0,00035	563.465	25.388.509	0,000	98,76%	25,93%	0	0	0
3	6 grupos (solo clúster)	6.839.078,77	7.858.011	7.304.318,18	5	0,000	0,00033	14.348	24.085.621	0,000	99,97%	82,84%	4,48%	77,24%	2,34%
4	6 grupos iguales	7.640.340,22	7.858.011	7.340.340,22	5	0,000	0,00035	39.865	28.113.819	0,000	99,95%	22,79%	0,00%	42,82%	0,00%
5	4 grupos iguales (cuartiles)	7.367.620,19	7.858.011	7.367.620,19	3	0,000	0,00035	39.057	21.745.675	0,000	99,94%	77%	2,24%	82,84%	4,48%

Fuente: elaboración propia a partir de los resultados EMP 2021

## Referencias bibliográficas del Apéndice 1.

- Anderberg, Michael R (1973). Cluster analysis for applications. Technical report, Office of the Assistant for Study Support Kirtland AFB NMEX.
- Green S. B. & Salkind N.J. (2008). Using SPSS for Window and Macintosh: Analyzing and understanding data (5th ed.). Upper Saddle River, NJ: Pearson Prentice Hall.
- Jardine, Nicholas and Sibson, Robin (1971). Mathematical taxonomy. London: John Wiley.
- Morris, Morris D. (1979). Measuring the Condition of the World's Poor: The Physical Quality of Life Index, New York, Pergamon.
- PNUD (1990). Informe de desarrollo humano 1990, Bogotá.
- PNUD (1992). Informe de Desarrollo Humano, Informe 1992, Bogotá, Tercer Mundo Editores
- Sokal, RR and Neath, PHAS (1963). ¿Principles of numerical taxonomy? Freeman. San Francisco–London.
- Sodeur. Wolfgang (2013). Empirische Verfahren zur Klassifikation, volume 42. Springer-Verlag.
- Varela Durán, Fausto (2017) Un estudio computacional de diferentes técnicas de clustering.
- Argandoña, A. (1990). El pensamiento económico de Milton Friedman. Universidad de Navarra. Recuperado en: <https://media.iese.edu/research/pdfs/DI-0193.pdf>.

## APÉNDICE 2. MODELO DE ESTIMACIÓN DE INGRESOS

Este apéndice explica el procedimiento metodológico empleado para definir y aplicar el modelo de estimación de ingreso a partir del SISBÉN IV. El procedimiento utilizado constituye una aproximación al uso de registros administrativos, para la estimación de coeficientes de regresión obtenidos de una fuente alternativa que permita “superar” el sesgo que se tiene en el SISBÉN IV respecto a los ingresos reportados.

De esta manera, para la identificación de esta fuente alternativa, se realizó un ejercicio de revisión de fuentes de información disponibles para ello, identificando que la Encuesta Multipropósito de Bogotá (EMB) 2021 sería el insumo idóneo, en la medida en la cual su contenido temático, su potencia estadística en términos de la muestra y su representatividad, permiten obtener resultados con gran precisión estadística a diferentes niveles de desagregación útiles para el análisis.

El propósito en este caso se traduce en definir un modelo estadístico a partir de la EMB y posterior a ello usar la información reportada en el SISBÉN IV para estimar el ingreso de la población allí registrada, de manera tal que, a partir de las condiciones reportadas en el SISBÉN IV se pueda aproximar el ingreso de las personas y hogares residentes en el Distrito.

Así las cosas, a continuación, se describe el conjunto de procesos, validaciones y revisiones realizadas para obtener el modelo que permitió obtener y reportar el ingreso estimado de la población objetivo.

Como primer y más importante proceso, fue necesario identificar la metodología estadística idónea para obtener la estimación del ingreso, pues aun cuando existe la posibilidad de implementar distintas metodologías, cada una de ellas debe ser evaluadas desde el punto de vista técnico, junto con sus restricciones, necesidades y viabilidades frente al conjunto de datos disponible en esta ocasión.

Para ello, a partir del escenario en el cual la variable de interés es el ingreso secundario, se excluye gran cantidad de metodologías que no involucran la posibilidad de estimar un parámetro con la condición del ingreso. Y, de otra parte, permite establecer que un modelo lineal generalizado resulta tener mayor capacidad de adaptación a la necesidad del estudio, al analizar como principal criterio de facilidad de uso, el que el valor esperado de la variable dependiente (Y) esta dada por un conjunto de parámetros e indicadores de interés (X) cumpliendo la siguiente ecuación:

$$E(Y|X) = \mu = g^{-1}(X\beta)$$

Esta ecuación se traduce en que el valor esperado del ingreso ( $Y$ ) es una combinación lineal del conjunto de parámetros e indicadores de interés ( $X$ ) con los parámetros desconocidos ( $\beta$ ) que serán estimados a través del modelo.

En este caso fue excluido cualquier tipo de método relacionado con el aprendizaje de máquina (inteligencia artificial) partiendo del hecho en el cual la validación de supuestos naturales asociados a la distribución del conjunto de datos, los supuestos asociados y la interpretación posterior, resultan ser imprescindibles y por ello, el uso de metodologías estadísticas frecuentistas tiene mayor peso de cara a los objetivos del estudio.

El segundo de los procesos se traduce en identificar aquel conjunto de variables, indicadores o parámetros de interés objeto de análisis, en la medida en la cual ambas fuentes de información (EMB – SISBÉN IV) disponen de grandes conjuntos de indicadores que en su mayoría pueden no ser de interés, y que estos a su vez, pueden no ser comparables entre las dos fuentes de información.

Esto resulta ser un criterio vital para avanzar el ejercicio, pues la principal restricción que se tiene a la hora de realizar la implementación de este se traduce en disponer de manera obligatoria de información con las mismas condiciones en ambas fuentes, lo cual se traduce no solo en tener las mismas variables, indicadores o parámetros de interés, sino que las escalas de medición y las categorías de estos sean equivalentes.

De esta manera, se reduce ampliamente el conjunto de variables, indicadores o parámetros de interés candidatos a hacer parte del modelo, y con ello que haya sido necesario proceder con ese segmento reducido de indicadores para luego ser contrastados con la literatura y seleccionar aquellos que efectivamente presuntivamente pueden apoyar el proceso de estimación del ingreso.

Una vez finalizado el proceso de revisión y selección de variables, se procede a realizar la implementación del modelo, con el fin de validar la posibilidad de obtener uno que cuente con los requerimientos técnicos necesarios y pueda realizarse el proceso de estimación.

En el documento de la MESEP<sup>36</sup> en la página 22 en el pie de página se menciona que el Modelo Hot Deck utilizado para imputación de ingresos en la construcción de la variable de ingreso per cápita para el cálculo de pobreza monetaria: "Este método se escogió después de evaluar cinco metodologías: imputación múltiple, máxima verosimilitud, DNP tradicional, DNP por estrato y «Hot Deck»".

Se menciona lo anterior dado que en el modelo de Hot Deck de la MESEP incluye variables como: Número de personas por hogar, Jefatura del hogar, Menores de edad, Personas Trabajando y Estrato que son la mayoría de las variables que se utilizaron para el modelo

En un primer ejercicio, se definió un modelo que incluyó todos los indicadores disponibles y preseleccionados, lo cual según la literatura se conoce como un modelo saturado, con el fin de identificar la participación conjunta que tiene cada indicador sobre el conjunto de datos de la EMB.

---

<sup>36</sup> DNP, DANE. Misión para el Empalme de las Series de Empleo, Pobreza y Desigualdad (MESEP). 2012



Posterior a ello, y con el fin de identificar el conjunto de indicadores que tuvieran mejor ajuste sobre el conjunto de datos, se empleó el método *stepwise* definido por (Efroymson, 1960), el cual se resume en implementar un proceso iterativo en el que a partir de la inclusión o exclusión de alguno(s) de los indicadores inicialmente definidos, se identifique el modelo que presente mejores cualidades frente a la calidad del mismo, criterios de bondad de ajuste, precisión en las estimaciones y su capacidad predictiva de cara a la implementación del ejercicio propuesto con el SISBÉN IV.

Basado en esto, el modelo que cumple las anteriores condiciones se define como:

$$\log(\text{Ingreso}) \sim \beta_0 + \beta_1 * X_1 + \beta_2 * X_2 + \beta_3 * X_3 + \beta_4 * X_4 + \beta_5 * X_5 + \beta_6 * X_6 + \beta_7 * X_7 + \beta_8 * X_8$$

Donde:

- $X_1$ : Número de personas en el hogar.
- $X_2$ : Sexo del jefe de hogar.
- $X_3$ : Presencia de menores de edad en el hogar.
- $X_4$ : Presencia de personas trabajando en el hogar.
- $X_5$ : Presencia de personas en condición de discapacidad en el hogar.
- $X_6$ : Tenencia de vivienda.
- $X_7$ : Nivel socioeconómico del hogar.
- $X_8$ : Nivel educativo del jefe de hogar.

A continuación, se reportan algunos de los criterios de bondad de ajustes obtenidos con este modelo:

*Tabla A2-1. Criterios de bondad de ajuste del modelo*

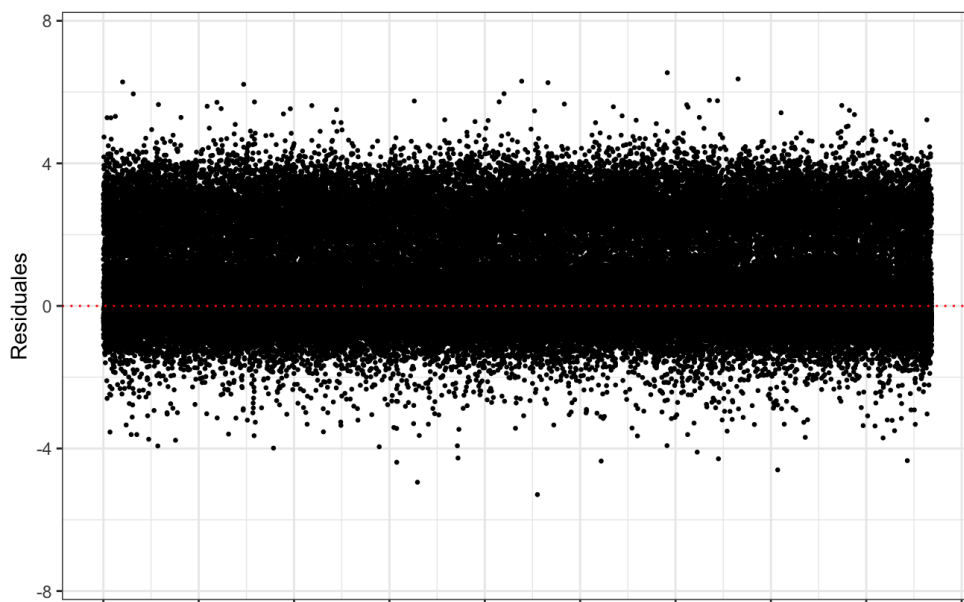
R2	Pseudo R2	AIC	BIC	Deviance
0,2366	0,2365	446.639,5	446808,2	869.629,8

*Fuente: elaboración propia a partir de los resultados de SISBÉN IV*

En la Tabla A2-1 se presentan el resultado de la medición de estos estadísticos, siendo estos a su vez, los que mejor comportamiento presentan al momento de comparar los diferentes modelos.

Por otra parte, el resultado del análisis de los residuales del modelo es favorable, en el sentido en el cual, no se evidencia ningún tipo de patrón o comportamiento no aleatorio que pueda afectar la confiabilidad de los valores predichos sobre el conjunto observaciones de la EMB. A continuación, un reporte visual de esta validación:

*Gráfico A2-1. Distribución de los residuales del modelo*



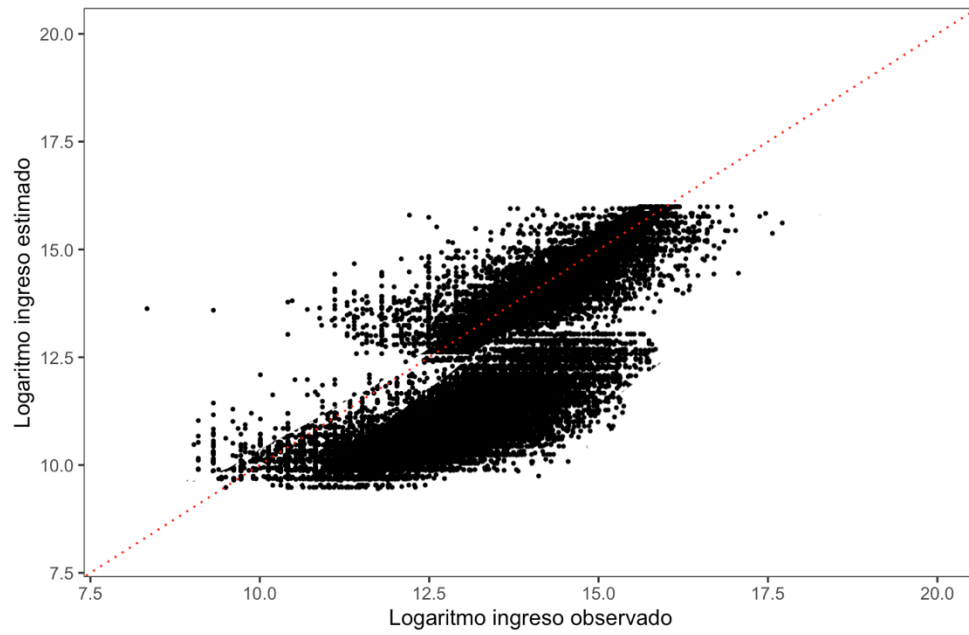
*Fuente: elaboración propia a partir de los resultados de SISBÉN IV*

Ahora bien, con el fin de confirmar explícitamente lo expuesto en el Gráfico A2-1, se validó con el test de Shapiro-Wilk referido en (Wilk, 1965), cuyo propósito se traduce en analizar la normalidad de los residuales, encontrando como resultado de ello, que no existe evidencia estadística suficiente (valor p superior a 0.05) para afirmar que los residuales no siguen una distribución normal.

Una vez validado el modelo, el último proceso a realizar se traduce en emplear los coeficientes de regresión definidos por el modelo y estimar el ingreso de los hogares registrados en el SISBÉN IV.

En el Gráfico A2-2 se presenta el contraste entre el ingreso estimado en la EMB 2021 vs su ingreso reportado en la EMB 2021, lo cual permite validar la consistencia de las estimaciones obtenidas y la capacidad predictiva que tiene el modelo de cara a su implementación con la información propia del SISBÉN IV.

Gráfico A2-2. Contraste logaritmo ingresos observados vs estimados



Fuente: elaboración propia a partir de los resultados de SISBÉN IV