



DIRECCIÓN GENERAL

GRUPO INTERNO DE TRABAJO EN POBREZA

NOTA METODOLÓGICA METODOLOGÍA DE EMPALME DE INDICADORES DEL AÑO 2021 A LAS SERIES DE POBREZA MONETARIA Y DESIGUALDAD DEL MARCO 2018

Septiembre de 2023

CONTENIDO

1. Introducción 3		
2. Justificación empalme series de indicadores de pobrez	za y desigualdad	4
3. Metodología de construcción de factor de empalme	8	
I. Partición geográfica donde se aplica la metodología	8	
II. Partición poblacional donde se aplica la metodología	8	
III. Factor de empalme 9		
IV. Uso del factor de empalme 9		
4. Evaluación del empalme realizado 9		
15		
5. Agradecimientos 15		

1. Introducción

El Departamento Administrativo Nacional de Estadística (DANE), en cumplimiento de su misión institucional y conforme a lo establecido en el CONPES 150 de 2012, es la entidad encargada de la estimación y oficialización de las cifras de pobreza. En la actualidad, existen dos mediciones complementarias: pobreza monetaria (y pobreza monetaria extrema) y pobreza multidimensional. Adicionalmente, el CONPES 150 también propuso la creación de un Comité de Expertos que tiene como principal objetivo garantizar la transparencia, comparabilidad y estabilidad metodológica de las mediciones de pobreza.

Con el propósito de generar cifras de pobreza monetaria, el DANE utiliza como fuente de información y análisis, la Gran Encuesta Integrada de Hogares (GEIH) que recoge información para la obtención de indicadores en materia de mercado laboral y pobreza monetaria en Colombia, que hasta el año 2021, utilizó el marco muestral del Censo Nacional de Población y Vivienda del año 2005. A partir de este año, se actualiza el marco muestral con el último Censo Nacional de Población y Vivienda del año 2018, un proceso que comienza en 2019 con la actualización de diseños muestrales, operativos y conceptuales. Adicionalmente, en la versión actualizada hay una introducción y ajuste de las preguntas del formulario, una nueva estratificación muestral de los municipios, ajuste de los tamaños de la muestra, nueva distribución de muestra y, por último, permitir una desagregación geográfica para las ciudades sin requerir sus áreas metropolitanas. Al respecto, en el año 2021 se realizan las dos encuestas una con el marco 2005 y otra con el marco 2018.

En este sentido, los resultados de los indicadores construidos con las variables de ingreso y mercado laboral de la encuesta en sus dos versiones del año 2021 (marco 2005 y marco 2018) fueron contrastados y analizados y se concluyó que presentan cambios de niveles en las series de los principales indicadores. Es por esto que, en un primer momento, se desarrolló un factor de empalme para los indicadores de mercado laboral que modifica los factores de expansión de la encuesta¹. Y, en segundo lugar, se construye un factor de empalme que garantiza la comparabilidad temporal (2021-2022) en los indicadores derivados de las variables de ingreso.

Este documento tiene como propósito describir la metodología empleada en la construcción del factor de empalme para las variables de ingreso de la encuesta 2021 marco 2005, con la que se calculó los

¹ DANE, Dirección de Metodología y Producción estadística (DIMPE). Gran Encuesta Integrada De Hogares Mercado Laboral. Nota Técnica: Empalme Series GEIH Marco 2005 - GEIH Marco 2018. Septiembre 2022.

datos oficiales de las incidencias de pobreza monetaria, pobreza monetaria extrema y coeficiente de Gini para el 2021.

2. Justificación empalme series de indicadores de pobreza y desigualdad

La Gran Encuesta Integrada de Hogares (GEIH) a partir del año 2021 actualiza el marco geoestadístico utilizando el Censo Nacional de Población y Vivienda de 2018, esto garantiza que las muestras y los datos recopilados reflejen con precisión la distribución geográfica y demográfica de la población. Adicionalmente se realiza una revisión e introducción de preguntas con el propósito de cumplir estándares de medición internacionales que puedan permitir la comparabilidad entre países y mejorar la precisión en la construcción de diferentes indicadores. Estos cambios pueden generar cambios de nivel en las series de los indicadores construidos a partir de la GEIH.

El DANE con el propósito de evaluar el impacto de los cambios metodológicos sobre las cifras estableció una estrategia de encuesta paralela en el año 2021. Donde simultáneamente se mide la encuesta utilizada hasta ese año 2021 con el marco de 2005 y la encuesta actualizada, que incorpora los cambios anteriormente nombrados, asociados al diseño muestral (CNPV 2018). Los resultados de esta comparación indican que, en los indicadores de pobreza y desigualdad de los grandes dominios geográficos, la variación porcentual entre marcos es mayor al error marginal². Esto implica que el cambio de nivel observado no se explica por la aleatoriedad de la muestra, sino que corresponde al cambio metodológico. Este resultado también se observa en algunas de las 23 ciudades principales (ver tabla 1; tabla 2 y tabla 3)

En este sentido, desarrollar la encuesta de esta manera permite identificar con precisión los cambios de nivel presentados en las series y se da la posibilidad de construir factores de empalme que permita cuantificar efecto y de esta forma permitir la comparabilidad entre las encuestas. En la siguiente tabla se muestra las cifras comparativas del marco censal 2005 y 2018 sin los tres primeros meses³ donde se presentan los resultados de pobreza monetaria, pobreza monetaria extrema y Gini y se adiciona el error marginal muestral con el fin de identificar las variaciones estadísticamente significativas.

² Para fines de comparación se estandarizó en ambas encuestas la Población en Edad de Trabajar (PET) y las proyecciones poblacionales. La PET pasó de ser de mayores de 10 o 12 años a una PET de mayores de 15 años² y las proyecciones de población² usadas son las creadas con el Censo Nacional de Población y Vivienda del año 2018.

³ No se toman los tres primeros meses debido a los altos niveles de tasas de no respuesta en las preguntas de ingreso.

Tabla 1. Tabla comparativa de pobreza monetaria para los marcos 2005 y 2018 del año 2021 sin los tres primeros meses.

	Pobreza Monetaria 2021			
Desagregación Geográfica	Marco 2005	Marco 2018	Variación valor absoluto	Error marginal
Total nacional	37,8%	39,5%	1,6%	0,5%
Cabecera	35,9%	36,8%	0,8%	0,5%
Centro				
poblado y	44,1%	48,4%	4,3%	1,2%
rural disperso				
Total				
Cabeceras 23				
ciudades más	34,6%	35,2%	0,6%	0,6%
Årea				
Metropolitana				
Resto urbano	38,4%	39,7%	1,3%	1,1%
Medellín	28,3%	31,1%	2,8%	1,2%
Barranquilla	36,5%	39,2%	2,7%	1,5%
Bogotá	33,2%	29,8%	3,4%	1,4%
Cartagena	36,5%	43,4%	6,9%	1,7%
Tunja	39,4%	45,9%	6,5%	1,9%
Manizales	27,7%	25,7%	2,0%	1,6%
Florencia	47,7%	50,7%	3,0%	1,7%
Popayán	43,1%	42,8%	0,3%*	1,8%
Valledupar	51,0%	49,3%	1,7%	1,6%
Montería	38,0%	36,9%	1,1%*	1,8%
Quibdó	60,9%	67,4%	6,6%	1,6%
Neiva	44,5%	47,7%	3,1%	2,0%
Riohacha	50,6%	49,9%	0,7%*	1,7%
Santa marta	47,6%	45,8%	1,8%	1,7%
Villavicencio	30,2%	27,1%	3,1%	1,6%
Pasto	37,2%	38,0%	0,8%*	2,0%
Cúcuta	46,0%	45,9%	0,1%*	1,7%
Armenia	38,6%	37,9%	0,7%*	1,7%
Pereira	32,4%	35,6%	3,2%	1,6%
Bucaramanga	36,2%	39,6%	3,3%	2,0%
Sincelejo	42,3%	47,5%	5,1%	1,7%
Ibagué	31,5%	45,9%	14,5%	1,7%
Cali	31,0%	33,2%	2,2%	1,6%

Fuente: DANE. GEIH 2021 meses abril-diciembre Nota: las variaciones con * marcan las diferencias que no superan al error marginal

Tabla 2. Tabla comparativa de pobreza extrema para los marcos 2005 y 2018 del año 2021 sin los tres primeros meses.

	Pobreza monetaria extrema 2021			
Desagregación Geográfica	Marco 2005	Marco 2018	Variación valor absoluto	Error marginal
Total nacional	11,6%	13,8%	2,2%	0,3%
Cabecera	9,7%	11,2%	1,6%	0,3%
Centro poblado y rural disperso	17,9%	22,4%	4,5%	0,9%
Total Cabeceras 23 ciudades más Área Metropolitana	8,3%	10,0%	1,8%	0,3%
Resto urbano	12,3%	13,4%	1,1%	0,7%
Medellín	5,6%	7,2%	1,6%	0,5%
Barranquilla	8,1%	11,8%	3,7%	0,7%
Bogotá	8,6%	7,5%	1,1%	0,7%
Cartagena	5,3%	11,7%	6,5%	0,6%
Tunja	7,7%	19,7%	12,0%	0,9%
Manizales	6,3%	5,1%	1,2%	0,7%
Florencia	12,4%	18,4%	6,0%	1,1%
Popayán	12,5%	17,1%	4,6%	1,0%
Valledupar	14,0%	14,6%	0,6%*	1,0%
Montería	7,3%	8,0%	0,7%*	0,9%
Quibdó	27,4%	34,8%	7,3%	1,4%
Neiva	10,0%	13,3%	3,2%	1,0%
Riohacha	23,3%	22,0%	1,3%	1,4%
Santa marta	16,8%	15,8%	1,0%	1,2%
Villavicencio	7,1%	6,2%	0,9%	0,8%
Pasto	6,7%	9,8%	3,1%	0,8%
Cúcuta	11,6%	16,2%	4,6%	1,0%
Armenia	8,4%	10,0%	1,6%	0,9%
Pereira	5,0%	10,2%	5,2%	0,7%
Bucaramanga	7,8%	9,8%	2,1%	1,0%
Sincelejo	6,7%	12,5%	5,8%	0,7%
Ibagué	4,2%	19,4%	15,2%	0,7%
Cali	8,9%	12,0%	3,1%	0,9%

Nota: las variaciones con * marcan las diferencias que no superan al error marginal

Tabla 3. Tabla comparativa de Gini para los marcos 2005 y 2018 del año 2021 sin los tres primeros meses.

	Gini 2021			
Desagregación Geográfica	Marco 2005	Marco 2018	Variación valor absoluto	Error marginal
Total nacional	0,522	0,565	0,043	0,007
Cabecera	0,508	0,549	0,042	0,007
Centro poblado y rural disperso	0,450	0,478	0,029	0,010
Total Cabeceras 23 ciudades más Área Metropolitana	0,509	0,552	0,043	0,008
Resto urbano	0,474	0,499	0,025	0,012
Medellín	0,498	0,529	0,031	0,005
Barranguilla	0,478	0,509	0,032	0,007
Bogotá	0,525	0,554	0,029	0,007
Cartagena	0,468	0,528	0,060	0,006
Tunja	0,466	0,564	0,099	0,009
Manizales	0,479	0,507	0,028	0,007
Florencia	0,487	0,532	0,044	0,011
Popayán	0,477	0,535	0,058	0,010
Valledupar	0,473	0,520	0,047	0,010
Montería	0,468	0,516	0,049	0,009
Quibdó	0,543	0,558	0,015	0,014
Neiva	0,464	0,505	0,041	0,010
Riohacha	0,549	0,547	0,001	0,014
Santa marta	0,479	0,524	0,045	0,012
Villavicencio	0,444	0,508	0,064	0,008
Pasto	0,487	0,519	0,032	0,008
Cúcuta	0,451	0,509	0,057	0,010
Armenia	0,455	0,490	0,035	0,009
Pereira	0,415	0,504	0,089	0,007
Bucaramanga	0,477	0,481	0,004	0,010
Sincelejo	0,470	0,504	0,034	0,007
Ibagué	0,445	0,558	0,113	0,007
Cali	0,495	0,540	0,045	0,009

Nota: las variaciones con * marcan las diferencias que no superan al error marginal

3. Metodología de construcción de factor de empalme

La metodología definida para la realización del empalme de las series de pobreza y desigualdad consistió en construir un factor de empalme para el año 2021 que permita comparar la distribución del ingreso per cápita del marco 2005 a la distribución del ingreso per cápita inducida por el marco 2018. Las metodología utilizada se elabora teniendo en cuenta los deciles de ingreso en cada uno de los dominios geográficos, ya que se encontró que era la presentaba una mayor precisión.

En el desarrollo del proceso se encontraron varios aspectos que afectaban la calidad del empalme:

- 1. Altas tasas de no respuesta en las preguntas de ingreso, particularmente en los primeros 3 meses de la encuesta de marco 2018.
- 2. Algunos dominios de desagregación geográfica con porcentajes mayores de ingreso per cápita en cero.

Por tales motivos se tomó la decisión de no considerar los primeros tres meses del año 2021 para construir el factor de empalme y trabajar los ceros del ingreso per cápita de forma independiente al proceso de empalme del ingreso mayor de cero.

A continuación, se presentan las etapas desarrolladas en la aplicación del método de empalme.

I. Partición geográfica donde se aplica la metodología

Tanto para el marco 2005 como para el marco 2018, se divide al país en 31 bloques geográficos consistentes con el diseño de la muestra actual. Los bloques están definidos en las 23 cabeceras capitales de departamento con sus respectivas áreas metropolitas; 4 regiones para el resto de las cabeceras y otras 4 regiones para los centros poblados y rural disperso. Las regiones son Atlántica, Pacífica, Oriental y Central.

II. Partición poblacional donde se aplica la metodología

Dentro de cada bloque definido en el ítem anterior se parte la población en dos grupos. Los que tienen ingreso cero y los que tienen ingreso mayor de cero. Con el segundo grupo se procede a identificar los deciles de la distribución de ingresos, lo cual genera una partición de diez grupos para la población con ingresos mayores de cero. En conclusión, se define 11 particiones de la población en cada bloque geográfico. Se aclara que en el proceso se toma como referencia el factor de expansión empalmado de mercado laboral.

III. Factor de empalme

El factor de empalme se calcula en cada partición definida previamente. Se espera que al multiplicar el factor de empalme con la variable de ingreso de marco 2005 se genere una nueva variable de ingreso ajustada con la cual se asegure una aproximación a las estructuras distribucionales de la variable de ingreso del marco 2018. La fórmula del factor de empalme es la siguiente:

$$Factor\ de\ empalme = \frac{Mediana\ del\ ingreso\ perc\'apita\ marco\ 2018}{Mediana\ del\ ingrso\ perc\'apita\ marco\ 2005}$$

IV. Uso del factor de empalme

Por último, el factor de empalme es puesto en la base de marco 2005 incluyendo los tres meses que se habían quitado en el proceso de creación del factor. Para hacerlo correctamente, previamente se particiona la población de la misma forma en que se particionó cuando se construyó el factor. Se crea una nueva variable de ingresos multiplicando el factor de empalme a la variable de ingresos per cápita de marco 2005, y con esta variable se procede a generar las estadísticas de pobreza y desigualdad.

4. Evaluación del empalme realizado

Para escoger la mejor opción de empalme, se verificó: que la metodología de empalme aplicada sobre la GEIH marco 2005 permita replicar los indicadores de interés (incidencia de la pobreza, de la indigencia y coeficiente de Gini) sin diferencias entre la GEIH marco 2005 y la GEIH marco 2018 estadísticamente significativas al 95% y que la metodología de empalme aplicada sobre la GEIH marco 2005 se aproxime a la distribución del ingreso de la GEIH marco 2018, de tal forma que, además de replicar los indicadores de interés , no exista dominancia estocástica entre la distribución del ingreso de la GEIH marco 2005 empalmada y la reportada por la GEIH Marco 2018.

Los resultados generados con el factor empalmado en los diferentes indicadores de pobreza fueron aceptables en la medida que las diferencias entre los indicadores generados después de aplicar el factor de empalme en el ingreso per cápita del marco 2005 disminuyeron con respecto al resultado del marco 2018. En muchos casos, la variación está contenida en el error marginal. Las siguientes tablas se muestran los resultados de pobreza, pobreza extrema y Gini.

Tabla 4. Tabla comparativa de pobreza monetaria para el marco 2005 empalmado y marco 2018 del año 2021 sin los tres primeros meses.

	Pobreza Monetaria 2021			
Desagregación Geográfica	Empalme 2005	Marco 2018	Variación valor absoluto	Error marginal
Total nacional	39,4%	39,5%	0,1%	0,5%
Cabecera	36,7%	36,8%	0,0%	0,5%
Centro poblado y rural disperso	48,3%	48,4%	0,1%	1,2%
Total Cabeceras 23 ciudades más Área Metropolitana	35,1%	35,2%	0,1%	0,6%
Resto urbano	39,7%	39,7%	0,0%	1,1%
Medellín	29,9%	31,1%	1,2%	1,2%
Barranquilla	39,4%	39,2%	0,2%	1,5%
Bogotá	30,6%	29,8%	0,8%	1,4%
Cartagena	42,3%	43,4%	1,1%	1,7%
Tunja	41,7%	45,9%	4,2%	1,9%
Manizales	26,4%	25,7%	0,7%	1,6%
Florencia	50,7%	50,7%	0,0%	1,7%
Popayán	41,8%	42,8%	1,0%	1,8%
Valledupar	49,7%	49,3%	0,4%	1,6%
Montería	37,2%	36,9%	0,3%	1,8%
Quibdó	67,6%	67,4%	0,2%	1,6%
Neiva	48,1%	47,7%	0,4%	2,0%
Riohacha	49,3%	49,9%	0,6%	1,7%
Santa marta	45,6%	45,8%	0,2%	1,7%
Villavicencio	26,9%	27,1%	0,2%	1,6%
Pasto	37,4%	38,0%	0,6%	2,0%
Cúcuta	45,9%	45,9%	0,0%	1,7%
Armenia	38,0%	37,9%	0,1%	1,7%
Pereira	35,0%	35,6%	0,6%	1,6%
Bucaramanga	40,1%	39,6%	0,5%	2,0%
Sincelejo	47,7%	47,5%	0,3%	1,7%
Ibagué	43,3%	45,9%	2,6%	1,7%
Cali	32,8%	33,2%	0,4%	1,6%

Tabla 5. Tabla comparativa de pobreza monetaria extrema para el marco 2005 empalmado y marco 2018 del año 2021 sin los tres primeros meses.

	Pobreza monetaria extrema 2021			
Desagregación Geográfica	Empalme 2005	Marco 2018	Variación valor absoluto	Error marginal
Total nacional	13,5%	13,8%	0,3%	0,3%
Cabecera	11,0%	11,2%	0,2%	0,3%
Centro poblado y rural disperso	21,6%	22,4%	0,9%	0,9%
Total Cabeceras 23 ciudades más Área Metropolitana	9,9%	10,0%	0,2%	0,3%
Resto urbano	13,2%	13,4%	0,2%	0,7%
Medellín	6,8%	7,2%	0,4%	0,5%
Barranquilla	10,8%	11,8%	1,0%	0,7%
Bogotá	8,1%	7,5%	0,6%	0,7%
Cartagena	10,2%	11,7%	1,5%	0,6%
Tunja	16,5%	19,7%	3,2%	0,9%
Manizales	5,4%	5,1%	0,3%	0,7%
Florencia	18,5%	18,4%	0,1%	1,1%
Popayán	17,0%	17,1%	0,1%	1,0%
Valledupar	15,2%	14,6%	0,6%	1,0%
Montería	9,0%	8,0%	1,0%	0,9%
Quibdó	34,1%	34,8%	0,7%	1,4%
Neiva	12,9%	13,3%	0,4%	1,0%
Riohacha	22,1%	22,0%	0,1%	1,4%
Santa marta	15,4%	15,8%	0,4%	1,2%
Villavicencio	6,0%	6,2%	0,2%	0,8%
Pasto	10,2%	9,8%	0,4%	0,8%
Cúcuta	16,2%	16,2%	0,0%	1,0%
Armenia	10,2%	10,0%	0,2%	0,9%
Pereira	10,4%	10,2%	0,2%	0,7%
Bucaramanga	9,6%	9,8%	0,3%	1,0%
Sincelejo	10,9%	12,5%	1,6%	0,7%
Ibagué	18,4%	19,4%	0,9%	0,7%
Cali	10,7%	12,0%	1,3%	0,9%

Tabla 6. Tabla comparativa de Gini para el marco 2005 empalmado y marco 2018 del año 2021 sin los tres primeros meses.

	Gini			
Desagregación Geográfica	Empalme 2005	Marco 2018	Variación valor absoluto	Error marginal
Total nacional	0,561	0,565	0,004	0,007
Cabecera	0,546	0,549	0,003	0,007
Centro poblado y rural disperso	0,460	0,478	0,018	0,010
Total Cabeceras 23 ciudades más Área Metropolitana	0,550	0,552	0,002	0,008
Resto urbano	0,491	0,499	0,008	0,012
Medellín	0,516	0,529	0,014	0,005
Barranquilla	0,508	0,509	0,001	0,007
Bogotá	0,562	0,554	0,007	0,007
Cartagena	0,513	0,528	0,015	0,006
Tunja	0,537	0,564	0,027	0,009
Manizales	0,507	0,507	0,000	0,007
Florencia	0,529	0,532	0,003	0,011
Popayán	0,524	0,535	0,010	0,010
Valledupar	0,515	0,520	0,004	0,010
Montería	0,514	0,516	0,002	0,009
Quibdó	0,558	0,558	0,000	0,014
Neiva	0,513	0,505	0,008	0,010
Riohacha	0,539	0,547	0,008	0,014
Santa marta	0,512	0,524	0,012	0,012
Villavicencio	0,496	0,508	0,012	0,008
Pasto	0,516	0,519	0,003	0,008
Cúcuta	0,492	0,509	0,017	0,010
Armenia	0,486	0,490	0,003	0,009
Pereira	0,490	0,504	0,014	0,007
Bucaramanga	0,481	0,481	0,001	0,010
Sincelejo	0,497	0,504	0,007	0,007
Ibagué	0,540	0,558	0,018	0,007
Cali	0,532	0,540	0,008	0,009

En lo que respecta a la distribución del ingreso se observa que la distribución de marco 2005 empalmado se aproxima satisfactoriamente a la distribución del marco 2018. Para ilustrar este fenómeno se tiene las gráficas de Pen's Parade, Lorenz y Lorenz generalizada para el total nacional.

Gráfico 1. Gráfica de Pen's Parade Total nacional año 2021 sin los tres primeros meses.

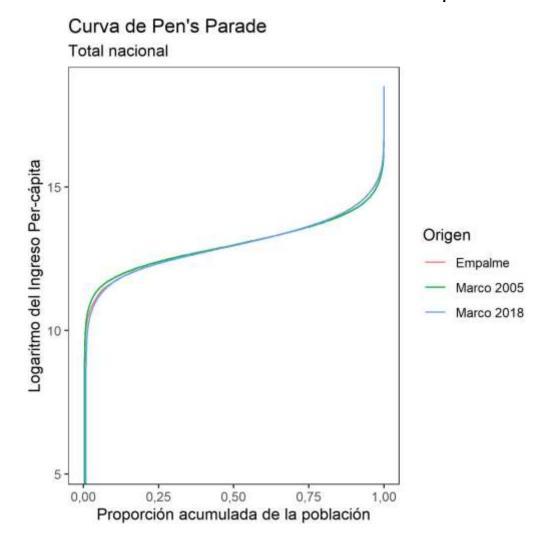


Gráfico 2. Gráfica Lorenz de Total nacional año 2021 sin los tres primeros meses.

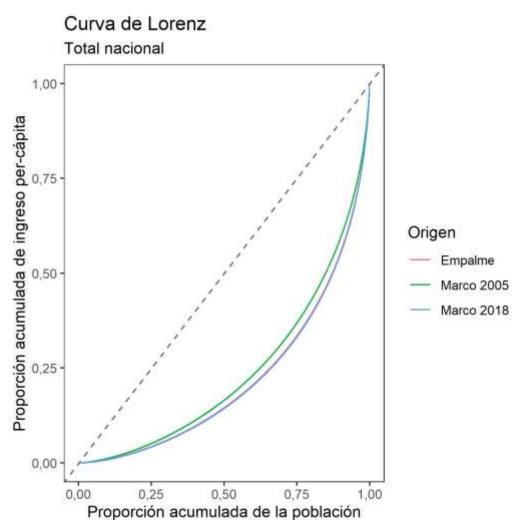


Gráfico 3. Gráfica Lorenz generalizada de Total nacional año 2021 sin los tres primeros meses.



5. Agradecimientos

Resulta importante reconocer el trabajo del Comité de Expertos en Pobreza, conformado por, Roberto Angulo, Adriana Camacho, Alejandra Corchuelo, Yadira Diaz, Renata Pardo, Mónica Pinilla, Juan Carlos Ramírez, Luis Alfredo sarmiento, Carlos Sepúlveda, María Dávalos (Banco Mundial), Jaime Urrego (PNUD), Sandra Álvarez, Guillermo Rivas y Rosa Katherine Rodríguez del equipo técnico del Departamento Nacional de Planeación y al equipo de Departamento de Prosperidad Social. Así mismo, fue fundamental el trabajo del Grupo Interno de Trabajo de Pobreza conformado por Olga Rocío Bermúdez, Francisco José Chaux, Ericson Gabriel Osorio y Sara Catalina Serrano. Finalmente, agradecer la labor de Carlos Arturo Ramírez Hernández en la coordinación del equipo y de Juan Sebastián Ordóñez del equipo de Registros Administrativos.