

Anexo A. Descripción técnica de la muestra

Diseño de la muestra

En Colombia, la muestra fue diseñada para que incluyera a todos los adultos no institucionalizados; es decir que excluye a las personas que habitan en cárceles, colegios, hospitales y bases militares. Es una muestra aleatoria estratificada. La estratificación asegura la inclusión de las áreas geográficas más importantes del país: Pacífica, Atlántica, Central, Oriental, los Antiguos Territorios Nacionales (atr) y Bogotá. La muestra fue subestratificada para incluir a las ciudades con más y con menos de 300.000 habitantes. Finalmente, la muestra fue subestratificada además en áreas urbanas y rurales.

La selección de la muestra fue multietápica. La primera etapa fue el municipio, luego el sector censal, luego la sección y finalmente la manzana, la vivienda y el hogar. Utilizamos un sistema de cuotas por sexo y edad para seleccionar al encuestado al interior de cada hogar.

Se entrevistaron 1491 informantes. Técnicamente, nuestro margen de error fue $\pm 2.54\%$. Esto quiere decir que si hiciéramos múltiples muestras en Colombia, 95% de ellas reflejarían las opiniones de la población con una precisión no inferior a $\pm 2.54\%$. Sin embargo, nuestra muestra fue estratificada y por conglomerados. Esto significa que, si bien la estratificación aumentó la precisión de la muestra, de alguna manera los conglomerados, que se hicieron para controlar costos del trabajo de campo, contribuyeron a reducir la precisión. Por supuesto, otros factores distintos al muestreo pueden reducir la precisión de la muestra, tales como la tasa de no respuesta, los errores al seleccionar al encuestado, la comprensión errónea de las preguntas, entre otros. Pero en términos de la técnica de muestreo, un intervalo de confianza de $\pm 2.54\%$ es muy bueno.

La Tabla A-1 es un resumen de los errores estándar y de los efectos de diseño para algunas variables e índices de la encuesta, para el año 2007. Los efectos de diseño (DEF) indican la eficiencia del diseño por conglomerados en comparación con el diseño aleatorio simple. Un DEF de 1 indica que las varianzas obtenidas en ambos diseños son las mismas, es decir que el diseño por conglomerados fue tan eficiente como un diseño aleatorio simple. Si el DEF es mayor que 1, significa que el diseño por conglomerados tuvo una varianza mayor que la producida por un diseño aleatorio simple. Y si el DEF es menor que 1, significa que la varianza del diseño por conglomerados es aun más pequeña que la producida por el diseño aleatorio.

Tabla A-1. Errores estándar y efectos de diseño para algunas variables (2007)

Variable	Media	Error est.	DEF
q2 (edad)	36.633	.323	.584
ed (nivel educativo)	9.360	.222	2.923
Wealth (índice de riqueza)	4.982	.127	4.519
PSA5 (respaldo al sistema)	57.987	.725	1.557
tol (tolerancia política)	54.898	.967	1.565
soct1r (evaluación de la economía del país)	41.010	.654	1.077
l1 (autoubicación ideológica)	6.120	.090	1.532
m1r (evaluación al gobierno)	64.705	.801	1.477
exc7r (percepción de corrupción)	74.600	1.006	1.397

estrato=región (estratopri)
upm=sector (colsect)

De acuerdo con la tabla anterior, el diseño por conglomerados para esta encuesta fue muy eficiente. En efecto, con la excepción de edad, educación y riqueza, todos los DEF estuvieron cerca de 1. Los errores estándar para la mayoría de variables también fueron muy moderados.

Descripción técnica del diseño muestral

Universo

El universo de la encuesta tiene una cobertura nacional de todos los adultos que habitan en las seis regiones del país: Bogotá, las regiones Atlántica, Pacífica, Central, Oriental y los Antiguos Territorios Nacionales. El universo también comprendió a todos los adultos que habitan en las áreas urbanas y rurales.

El universo se dividió en dos sectores: uno con las ciudades con más de 300.000 habitantes y otro con las ciudades con menos de 300.000 habitantes.

Población

La muestra se circunscribió a todos los adultos no institucionalizados; es decir que excluye a las personas que habitan en cárceles, colegios, hospitales y bases militares. Los hogares privados en estas áreas sí fueron contemplados.

Unidad de selección final

Dado que el cuestionario incluye preguntas no sólo referentes al encuestado sino también a otros miembros del hogar, la unidad de observación estadística seleccionada fue el hogar. El encuestado sólo podía habitar en un hogar.

Dado que cada hogar pertenece a una vivienda, a veces compartida con otros hogares relativamente estables en el tiempo, cada vivienda fue seleccionada como la unidad de selección final.

Método de muestreo

Escogimos el muestreo probabilístico, estratificado, multietápico, con selección aleatoria de las unidades en cada etapa. Primero, la muestra fue estratificada por tamaño de municipio (ciudades con más y menos de 300.000 habitantes), luego por regiones y luego por áreas (rural y urbana).

Es un muestreo multietápico puesto que al interior de las áreas urbanas se empezó con unidades primarias de muestreo (sectores), luego con unidades secundarias (secciones), luego con unidades terciarias (manzanas) y con unidades finales de muestreo (conglomerados de viviendas) de 6 a 8 en áreas urbanas y de 10 a 12 en áreas rurales. En cada vivienda el encuestador seleccionó un solo hogar como unidad de observación.

El informante fue seleccionado de acuerdo con cuotas de edad y sexo. En efecto, en cada manzana el encuestador tenía que incluir al menos un hombre y una mujer de los siguientes rangos de edad:

- 18 a 27 años
- 28 a 40 años
- Más de 40 años

A cada encuestador se le asignó una manzana específica. Una vez en la manzana, el encuestador hacía un recuento de las primeras 20 viviendas que encontraba. Tenía la instrucción de aplicar un mínimo de 8 encuestas de las 20 viviendas listadas, procurando un balance de las cuotas de sexo y edad.

El método de selección se eligió de acuerdo con las siguientes consideraciones:

Se requerían muestras representativas en los siguientes niveles:

- Nacional
- Estratos de primera etapa:
 - o Ciudades con más de 300.000 habitantes
 - o Ciudades con menos de 300.000 habitantes
- Estratos de segunda etapa:
 - o Bogotá
 - o Región Atlántica
 - o Región Pacífica
 - o Región Oriental
 - o Región Central
 - o Antiguos Territorios Nacionales

- Estratos de tercera etapa:
 - o Área urbana
 - o Área rural

- Dominios:
 - o Ciudades con más de 300.000 habitantes (de inclusión forzosa)
 - o Ciudades con menos de 300.000 habitantes

Para cada etapa, se calcularon márgenes de error que correspondieran a los estándares mínimos de calidad.

Igualmente, se trabajó con el mejor y más actualizado marco muestral disponible para cada municipio (censo poblacional, cartografía, listados de viviendas actualizados, entre otros)

Marco muestral

El marco muestral está constituido por el inventario actualizado de cartografía y viviendas obtenidos a partir del Censo de 1993. El Centro Nacional de Consultoría obtuvo las versiones 2003 del Departamento Nacional de Estadística DANE.

Cálculos por estratos

La muestra está constituida por 193 puntos muestrales: 160 urbanos y 33 rurales, distribuidos en 53 municipios de 26 de los 32 departamentos de Colombia.

Tamaños de muestra, intervalos de confianza y márgenes de error

El intervalo de confianza anticipado para la muestra nacional fue de 95% con un margen de error de 2.6%, asumiendo una proporción 50/50 en variables dicotómicas.

Los márgenes de error para un intervalo de 95% de confianza se indican en la Tabla A-2:

Tabla A-2. Tamaño de la muestra y márgenes de error (nivel de confianza del 95%)

Estratos Regiones	Tamaño de muestra	Margen de error M.A.S.	% M.P.C.
Atlántica	323	5.45	5.58
Bogotá	231	6.45	6.59
Central	358	5.18	5.29
Oriental	274	5.92	6.05
Pacífica	251	6.19	6.32
Antiguos Territorios Nacionales	54	13.34	13.63
Áreas			
Urbana	1107	2.96	3.01
Rural	384	5.00	5.11
Total País	1491	2.54	2.6

Personal que trabajó en esta encuesta

El CNC involucró a sus cinco sedes (Bogotá, Cali, Medellín, Barranquilla y Bucaramanga) para asegurar una calidad óptima en el mínimo tiempo posible. Debido a la actual situación de orden público del país, se nos aconsejó permanecer el mínimo tiempo posible en la mayoría de áreas visitadas, lo cual complicó las operaciones.

Debido a la complejidad del formulario, utilizamos nuestros encuestadores más experimentados, muchos de ellos con más de 15 años de experiencia.

El CNC involucró a un total de 99 miembros en el equipo, distribuidos como se ve en la Tabla A-3:

Tabla A-3. Personal del CNC que participó en el estudio

Actividad	Total de personal
Coordinadores de campo	5
Supervisores	15
Encuestadores	40
Supervisores de calidad en campo	10
Codificadores	7
Digitadores	7
Verificadores de datos	7
Subtotal campo y digitación	91
Personal directivo y profesional	5
Personal administrativo	3
Total personal del equipo	99

Tabla A-4. Tamaño y distribución de la muestra por estrato (región)

Región	Urbana	Rural	Total
Atlántica			
> 300,000 habitantes	127		127
< 300,000 habitantes	112	84	196
Total Atlantica	239	84	323
Bogotá			
> 300,000 habitantes	231		231
Total Bogota	231		231
Central			
> 300,000 habitantes	150		150
< 300,000 habitantes	112	96	208
Total Central	262	96	358
Oriental			
> 300,000 habitantes	66		66
< 300,000 habitantes	112	96	208
Total Oriental	178	96	274
Pacífica			
> 300,000 habitantes	91		91
< 300,000 habitantes	88	84	160
Total Pacifica	179	72	251
Antiguos Territorios Nacionales			
< 300,000 habitantes	18	36	54
Total Antiguos Territorios Nacionales	18	36	54
Total	1,107	384	1,491

Tabla A-5. Encuestados y unidades primarias de muestreo (UPM) por estrato (región)

Región	Muestra	UPM
Atlántica	323	32
Bogotá	231	34
Central	358	39
Oriental	274	26
Pacífica	251	26
Antiguos Territorios Nacionales	54	6
Total	1,491	163

Comentarios finales sobre el trabajo de campo

Sobre el cuestionario: el cuestionario fue más largo que las anteriores versiones de LAPOP, 55 minutos en promedio. Sin embargo, en general los informantes estuvieron dispuestos a contestar las encuestas y tuvimos muy pocas entrevistas interrumpidas.

Acerca del trabajo de campo: Aparte de algunos incidentes en ciertos municipios con serios problemas de orden público incidentes, a pesar de que algunos informantes estaban ubicados en áreas con fuerte presencia de grupos armados ilegales, no hubo reportes de ningún tipo de presión sobre los informantes para inducir las respuestas. Por el contrario, los encuestadores enfatizaron la libertad de opinión de quienes aceptaron ser parte de este estudio.

Como todos los años, el CNC quisiera extender su gratitud a cada uno de los miembros del equipo de este estudio, especialmente a los valientes hombres y mujeres que desafiaron las normas de seguridad y asumieron muchos riesgos para lograr un trabajo muy bien hecho.