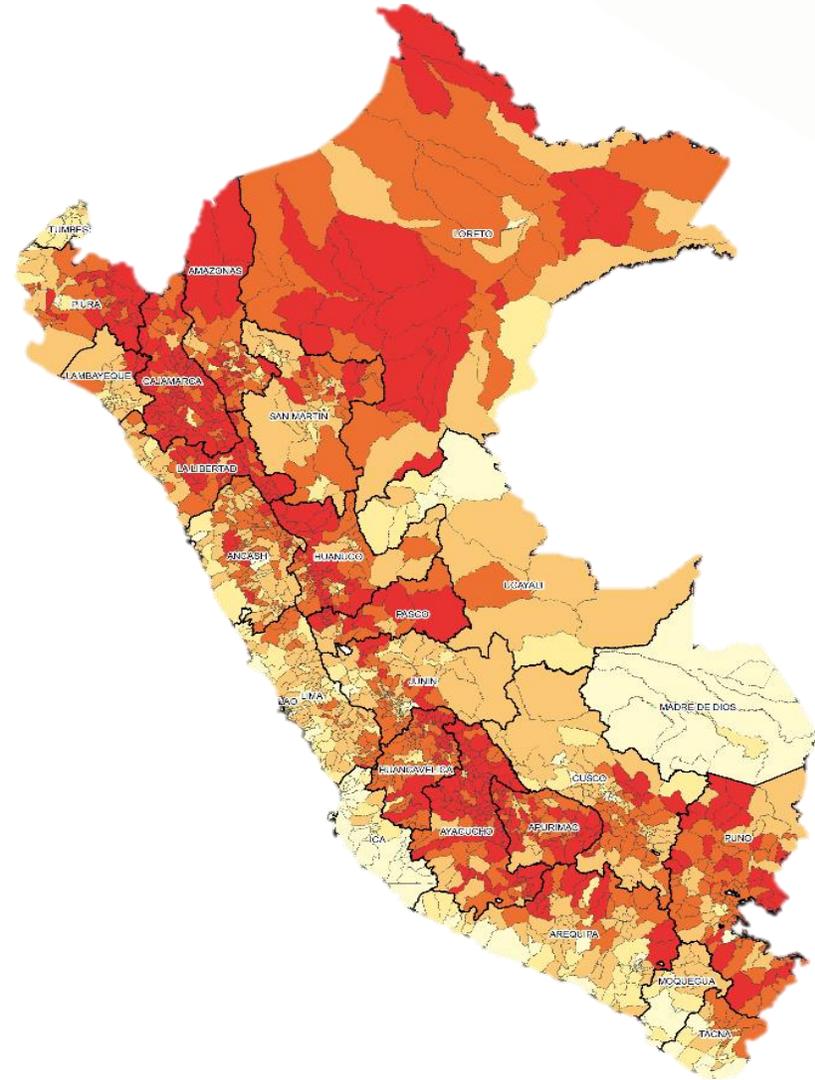


Seminario: “Integración de fuentes de datos y estadísticas oficiales”

Experiencia PERUANA en la medición de la pobreza monetaria a nivel distrital

Cirila Gutiérrez Espino
Directora Técnica de Demografía e Indicadores Sociales
INEI-Perú



CONTENIDO



OBJETIVOS



ANTECEDENTES Y NORMATIVIDAD PARA USO DE POLÍTICAS PÚBLICAS



DEFINICIÓN DE POBREZA



METODOLOGÍA Y FUENTES DE INFORMACIÓN

- Metodología
- Fuentes de información
- Preparación de base de datos



ETAPAS DEL PROCESO

- Procedimiento de enlace de bases de datos
- Preparación de los datos
- Modelamiento - proceso
- Validación
- Agrupación de distritos
- Limitaciones



RESULTADOS

- Estimación de la pobreza total por provincias y distritos.

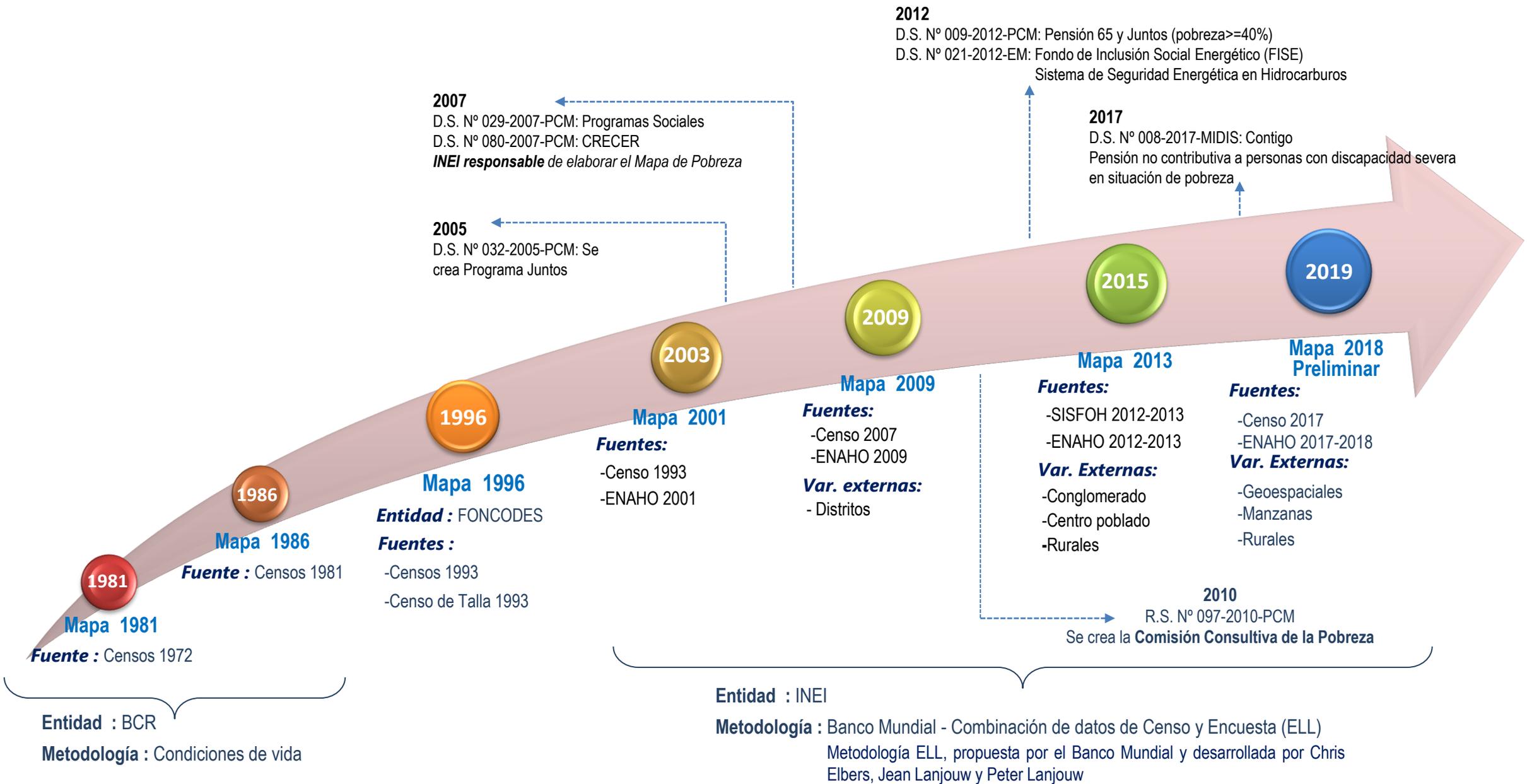
• OBJETIVO GENERAL

- ❑ Estimar la incidencia de la pobreza monetaria a nivel de áreas menores (**provincia, distrito e infra-distrito**), que permita identificar y priorizar las zonas más pobres del país, para **focalizar los programas sociales y diseñar las políticas públicas** con criterios de descentralización y articulación territorial.

• OBJETIVOS ESPECÍFICOS

- ❑ Identificar prioridades para la asignación de recursos públicos para mejorar la focalización del gasto público en zonas geográficas que presentan una mayor incidencia de la pobreza.
- ❑ Agrupar distritos de **similares características** para identificar sus carencias e implementar programas sociales de intervención.
- ❑ Desagregar distritos que presentan **alta heterogeneidad** para diferenciar las áreas con mayor pobreza.

ANTECEDENTES/NORMATIVIDAD PARA USO DE POLÍTICAS PÚBLICAS



Definición de pobreza

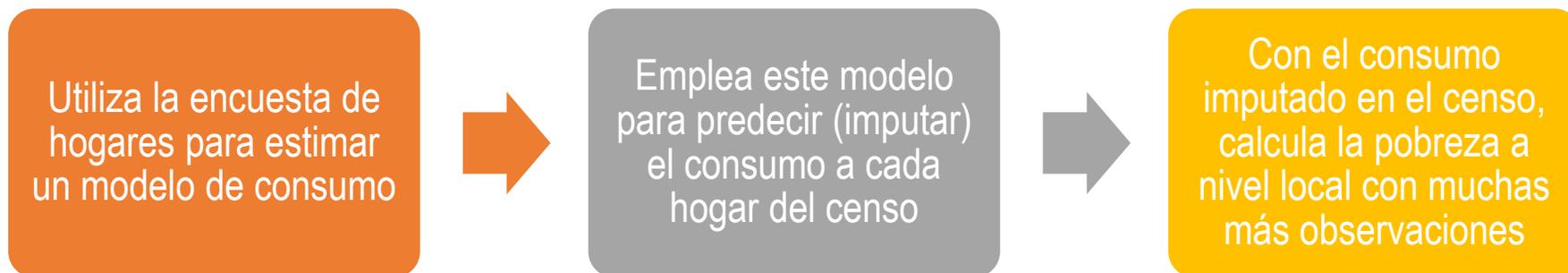
- **El Mapa de pobreza a nivel distrital**, considera **el enfoque monetario de la pobreza**. Se considera pobre a las personas que residen en hogares cuyo gasto total per cápita mensual está por debajo del valor de una canasta de *productos y servicios (líneas de pobreza) que permite satisfacer un conjunto de necesidades mínimas de bienes y servicios*". Por lo tanto, la pobreza monetaria mide la posibilidad de adquirir dicha canasta mínima de productos y servicios, pero no necesariamente su consumo efectivo.
- **Población en pobreza extrema** a "*personas que residen en hogares cuyo gasto per cápita mensual está por debajo del valor de la canasta de alimentos que permite satisfacer los requerimientos nutricionales mínimos (calorías) de los miembros del hogar*".

METODOLOGÍA Y FUENTES DE INFORMACIÓN

Metodología

- Metodología desarrollada por Chris Elbers, Jean Lanjouw y Peter Lanjouw del Banco Mundial.
- En la aplicación de la metodología, se combinaron los resultados de la Encuesta Nacional de Hogares (ENAH) 2017-2018 y del Censo de Población y Vivienda (CPV) 2017. Con el fin de hacer comparables la encuesta y el Censo, se armonizó el concepto de miembros del hogar aplicando la metodología empleada en la ENAH, y se ajustó la población total (censada más omitida) por grupos de edad, sexo, área de residencia y departamento.

Pasos:



Variables comunes y con distribuciones similares en censo y encuesta

Hogares con mismas características tienen diferente consumo por razones no observadas a nivel geográfico

$$\ln y_{ch} = X'_{ch}\beta + \eta_c + \varepsilon_{ch}$$

Hogares con mismas características tienen diferente consumo por razones no observadas a nivel de hogar

Fuentes de información

- Además, se recolectaron y prepararon 22 bases de datos que se utilizaron en la estimación de los modelos predictivos, de los cuales 20 bases es de otras fuentes externas: IV Censo Nacional Agropecuario (CENAGRO) 2012, Programa Nacional CUNAMAS 2018, Registro Nacional de Usuarios de Programas Sociales (RNU) 2017, Plan de Salud Escolar (PSE) 2016, Programa Nacional de Becas y Crédito Educativo (PRONABEC) 2012-2018, Registro Nacional de Municipalidades (RENAMU) 2018, Seguro Integral de Salud (SIS) 2017, Organismo Supervisor de la Inversión en Energía y Minería (OSINERMIN) 2017, entre otros.
- La disponibilidad de las diversas fuentes de información permitió una mejor aplicación de la metodología de elaboración del mapa de pobreza, en la medida que se requiere **igualdad en la distribución de las variables predictivas del gasto** según las fuentes censales y la encuesta de hogares, reduciendo los efectos fijos a distintos niveles de desagregación geográfica, así como generar un **sistema de bases de datos enlazados geográficamente al Censo 2017**, que constituye una herramienta útil para futuras investigaciones.

Preparación de bases de datos

- Diferentes bases de datos recolectadas, se construyeron un conjunto de variables que conformaron la base de datos final para la estimación de los modelos predictivos del gasto per cápita de los hogares.

Se diferenció entre **variables cualitativas** (como por ejemplo, sexo, estado civil, grado de instrucción) y **variables cuantitativas** (como por ejemplo, total de trabajadores, número de micro-pequeñas empresas).

Las **variables válidas** si los datos faltantes no superaron el 30% de los casos en las variables cualitativas, y el 40% en las variables cuantitativas.

ETAPAS DEL PROCESO

ETAPAS DEL PROCESO

1 ENLACE DE LAS BASES

- Recolectar fuentes de datos.
- Enlazar las bases de población e instituciones al Censo 2017.

2 PREPARACIÓN DE LOS DATOS

- Verificar la igualdad de las definiciones en el Censo y encuesta.
- Armonizar las variables.
- Ajustar el factor de expansión de la encuesta a la población total.
- Seleccionar variables usando tests estadísticos.
- Eliminar conglomerados con menos de 3 hogares.

3 MODELAMIENTO

- Obtener los modelos de consumo a nivel de regiones.
- Obtener los estadísticos de bondad de ajuste.

4 VALIDACIÓN

- Realizar la validación interna de los modelos mediante la sensibilidad de las estimaciones la pobreza total por regiones, dominio geográfico, conos, y en la submuestra; así como del gasto.

5 AGRUPACIÓN Y DESAGREGACIÓN

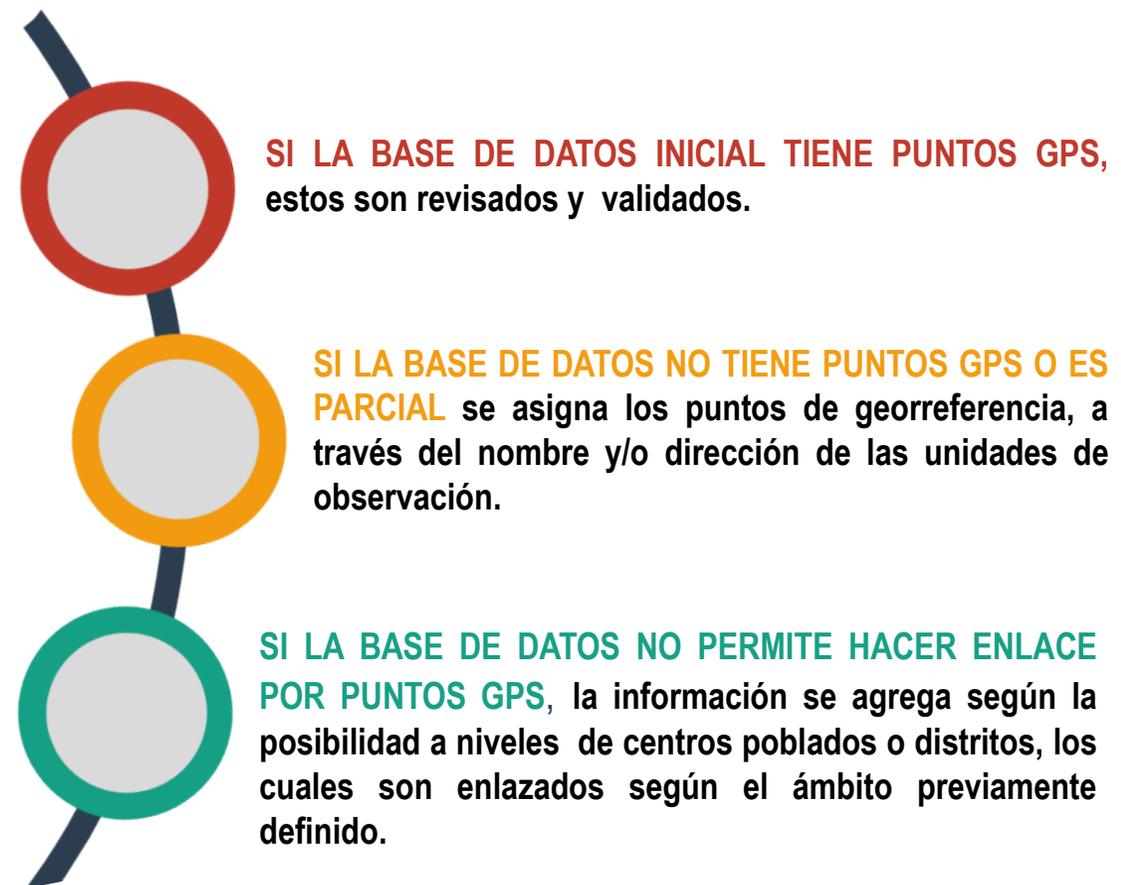
- Agrupar distritos con poblaciones pequeñas menor a mil hogares y que tengan poca precisión en la estimación de la pobreza ($CV \geq 25\%$).
- Desagregar distritos con población mayor a 20 mil hogares y que presenten heterogeneidad dentro de ellos.

1. Procedimiento de enlace de bases de datos

BASE DE POBLACIÓN



BASE DE INSTITUCIONES



Asignación agregada según ámbito



2. Preparación de los datos

Justificación

Las variables comunes entre el censo y la encuesta deben tener la misma definición y construcción como el tamaño del hogar, hogares con algún miembro que cuenta con seguro de salud, número de ocupados, entre otros.

Criterio

- Unir la ENAHO 2017-2018.
- Ajustar el factor de expansión a la población total por departamento, área de residencia, sexo y grupo de edad.

Conceptualización y estandarización de las variables entre el Censo 2017 y ENAHO 2017-2018

Material predominante en las paredes exteriores

CPV 2017

3. EN LA VIVIENDA, ¿EL MATERIAL DE CONSTRUCCIÓN PREDOMINANTE EN LAS PAREDES EXTERIORES ES DE:

(Lea cada alternativa y rellene solo un óvalo ●)

1 Ladrillo o bloque de cemento?

2 Piedra o sillar con cal o cemento?

3 Adobe?

4 Tapia?

5 Quincha (caña con barro)?

6 Piedra con barro?

7 Madera (pona, tornillo, etc.)?

8 Triplay / calamina / estera?

9 Otro material?

ENAHO 2017

102A.EL MATERIAL PREDOMINANTE EN LAS PAREDES EXTERIORES ES:

¿Ladrillo o bloque de cemento?..... 1

¿Piedra o sillar con cal o cemento?..... 2

¿Adobe?..... 3

¿Tapia?..... 4

¿Quincha (caña con barro)?..... 5

¿Piedra con barro?..... 6

¿Madera? "Pona, Tornillo"..... 7

¿Estera? "Triplay, alamina"..... 8

¿Otro material?..... 9

(Especifique)

ENAHO 2018

102A.EL MATERIAL PREDOMINANTE EN LAS PAREDES EXTERIORES ES:

¿Ladrillo o bloque de cemento?..... 1

¿Piedra o sillar con cal o cemento?..... 2

¿Adobe?..... 3

¿Tapia?..... 4

¿Quincha (caña con barro)?..... 5

¿Piedra con barro?..... 6

¿Madera (pona, tornillo, etc.)?..... 7

¿Triplay/calamina/estera?..... 8

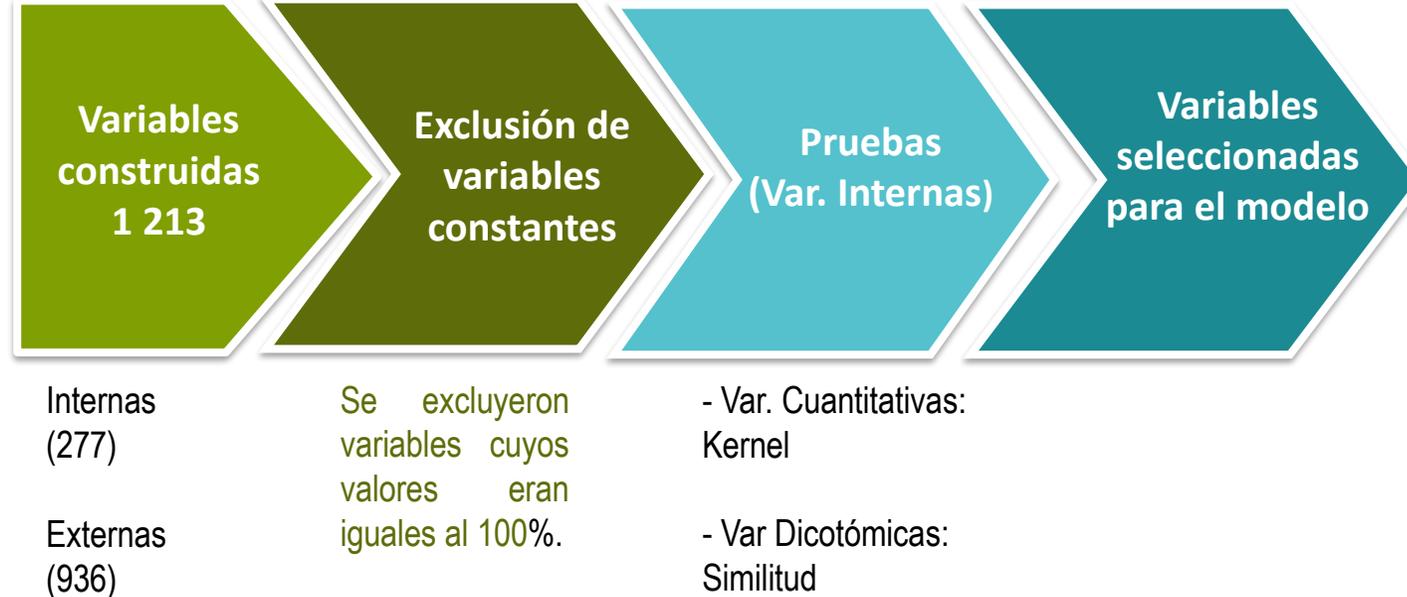
¿Otro material?..... 9

(Especifique)

3 Modelamiento - proceso

Se seleccionaron los modelos predictivos del gasto per cápita del hogar que presentaron las bondades de ajuste más elevadas y bajos ratios de varianza. La bondad del ajuste de la regresión, medida a través de los coeficientes de determinación ajustado (R^2) que indican cuánto de la variación del gasto es explicada por el modelo seleccionado, varía entre 0 y 1, siendo mejor el ajuste cuanto más se aproxima a 1.

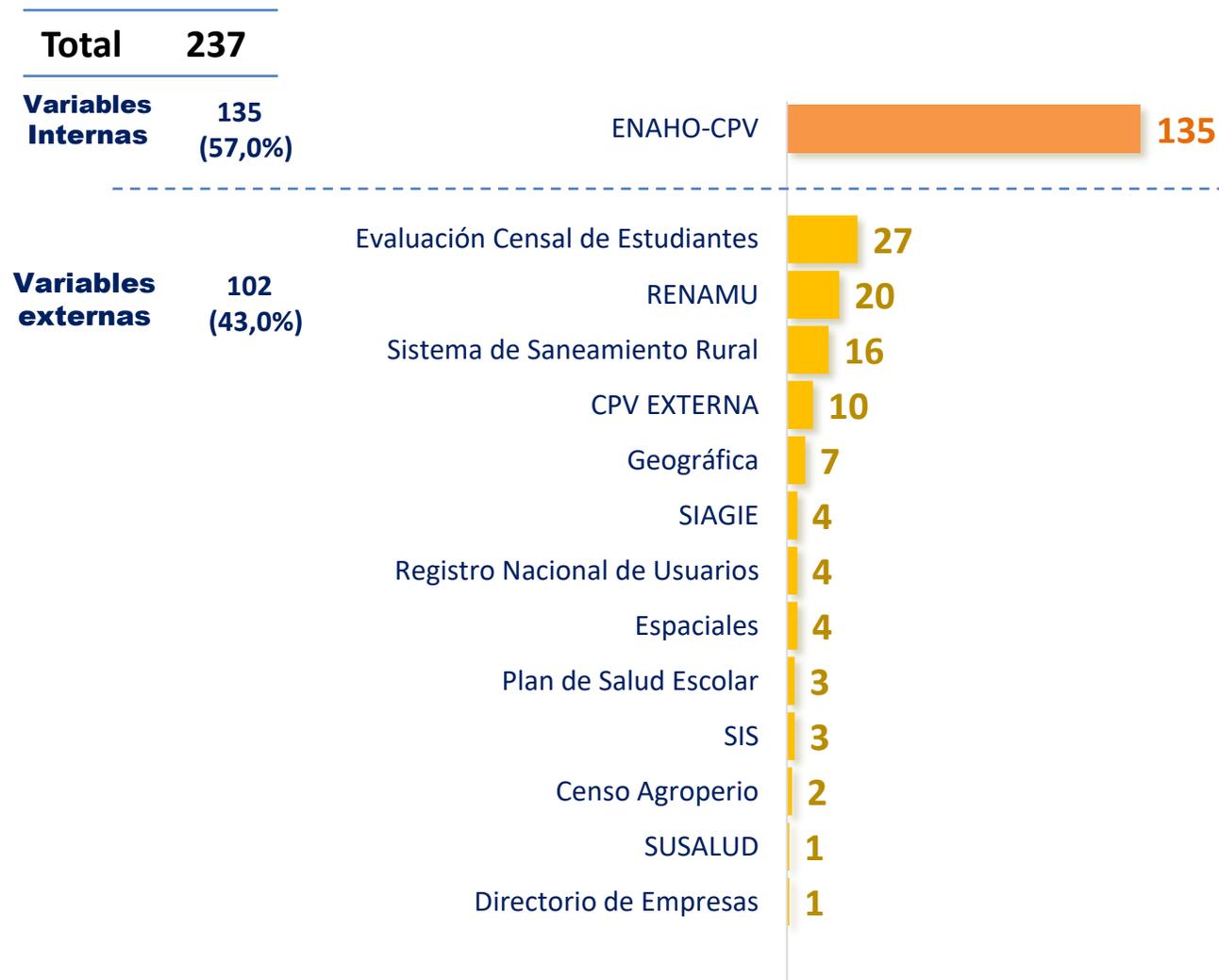
Proceso



Además de seleccionar variables definidas de manera estrictamente comparable, se examinó la igualdad en la distribución de las variables predictivas presentes tanto en la ENAHO 2017-2018 como en el Censo 2017. Se estimaron modelos econométricos con varios cortes geográficos.

3.1 Selección del modelo – Variables y bondad de ajuste

RESUMEN DE LAS VARIABLES EN EL MODELO



Ámbito	Número de variables en el modelo	R2 Ajustado	RMSE	Raíz cuadrada de las observaciones
TOTAL	237			
Loreto	23	72.4	0.370	54
Cusco	17	68.6	0.326	50
Provincia de Lima 1/	22	67.9	0.334	77
Huánuco	25	66.3	0.360	50
Junín	15	65.7	0.357	55
San Martín	18	64.1	0.369	51
La Libertad	28	63.5	0.394	56
Amazonas	23	63.1	0.377	49
Prov. Const. del Callao	14	61.7	0.315	45
Moquegua	19	61.7	0.353	44
Cajamarca	28	60.8	0.404	54
Piura	25	60.5	0.369	57
Región Lima 2/	23	60.0	0.319	53
Ucayali	17	58.9	0.322	48
Apurímac	18	57.6	0.313	43
Ayacucho	27	57.4	0.368	48
Ica	12	56.6	0.269	55
Huancavelica	26	55.1	0.335	45
Tumbes	17	53.0	0.331	41
Áncash	19	52.0	0.394	57
Pasco	16	50.1	0.368	46
Tacna	12	47.9	0.422	52
Arequipa	26	45.2	0.401	62
Madre de Dios	10	45.2	0.389	40
Lambayeque	14	42.3	0.421	53
Puno	26	41.0	0.429	54

1/ Comprende los 43 distritos que conforman la provincia de Lima.

2/ Comprende las Provincias de Barranca, Cajatambo, Canta, Cañete, Huaral, Huarochirí, Huaura, Oyón y Yauyos.

Fuente: INEI – Mapa de Pobreza Provincial y distrital, 2018

4. Validación

Esta etapa de la metodología tiene como finalidad verificar y garantizar la robustez de las estimaciones del gasto del hogar y de los resultados de pobreza monetaria. Se realizaron dos tipos de validaciones: interna y externa.

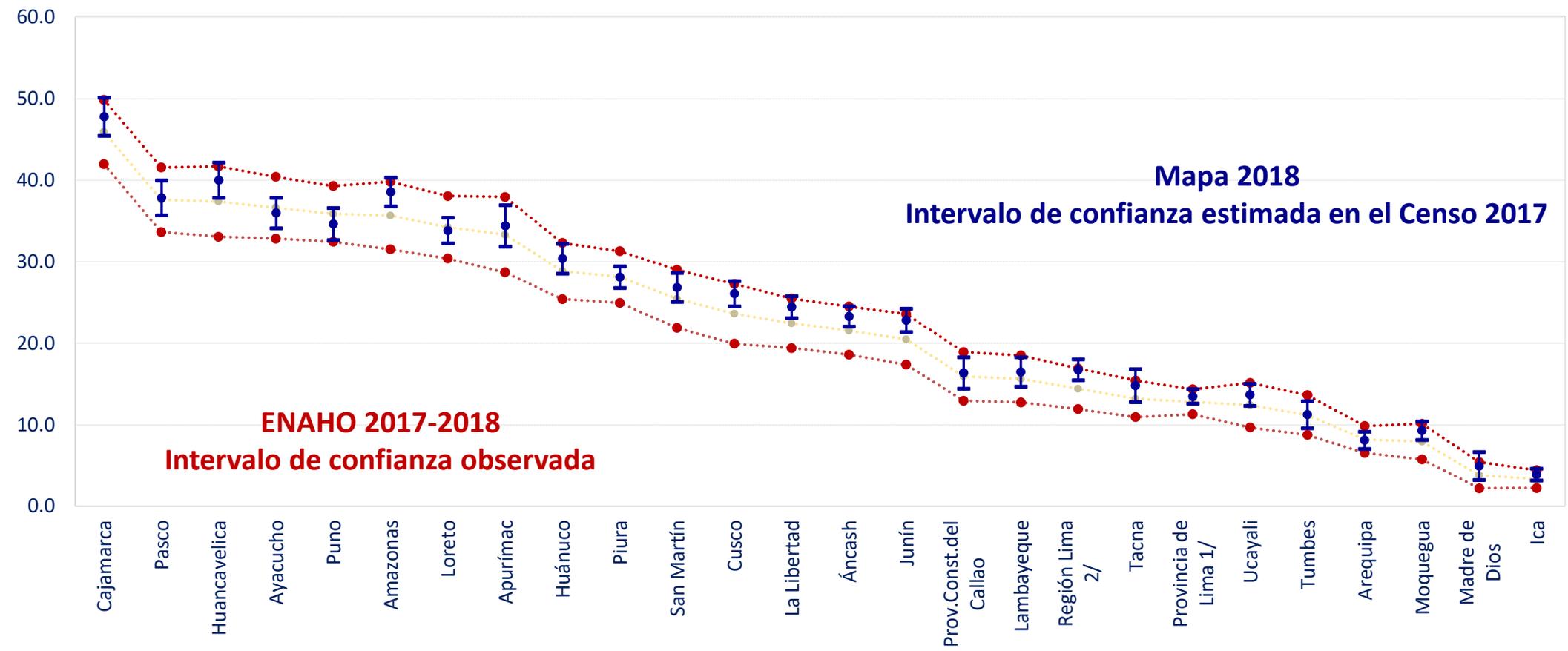
Para la validación interna se realizó el análisis de sensibilidad para evaluar la capacidad predictiva de los modelos del gasto.

- Se compararon las estimaciones de los modelos del gasto con las estimaciones directas de la ENAHO 2017-2018 por departamento.
- Se analizó la sensibilidad de las estimaciones del gasto en los modelos por departamento.

Validación interna – Sensibilidad por corte geográfico

INCIDENCIA DE LA POBREZA ESTIMADA CON EL CENSO 2017 E INCIDENCIA OBSERVADA CON LA ENAHO POR DEPARTAMENTOS, 2017-2018

Intervalos de confianza al 95% de la incidencia de pobreza monetaria obtenidas con la ENAHO 2017-2018 y de la estimación de la incidencia de pobreza monetaria obtenida con el Censo 2017. Se observa que todas las estimaciones de los modelos del gasto per cápita del hogar obtenidas con el Censo 2017 se encuentran dentro de los intervalos de confianza obtenidos con la encuesta, concluyendo que se ha obtenido una buena precisión en las estimaciones de la pobreza monetaria en todos los departamentos

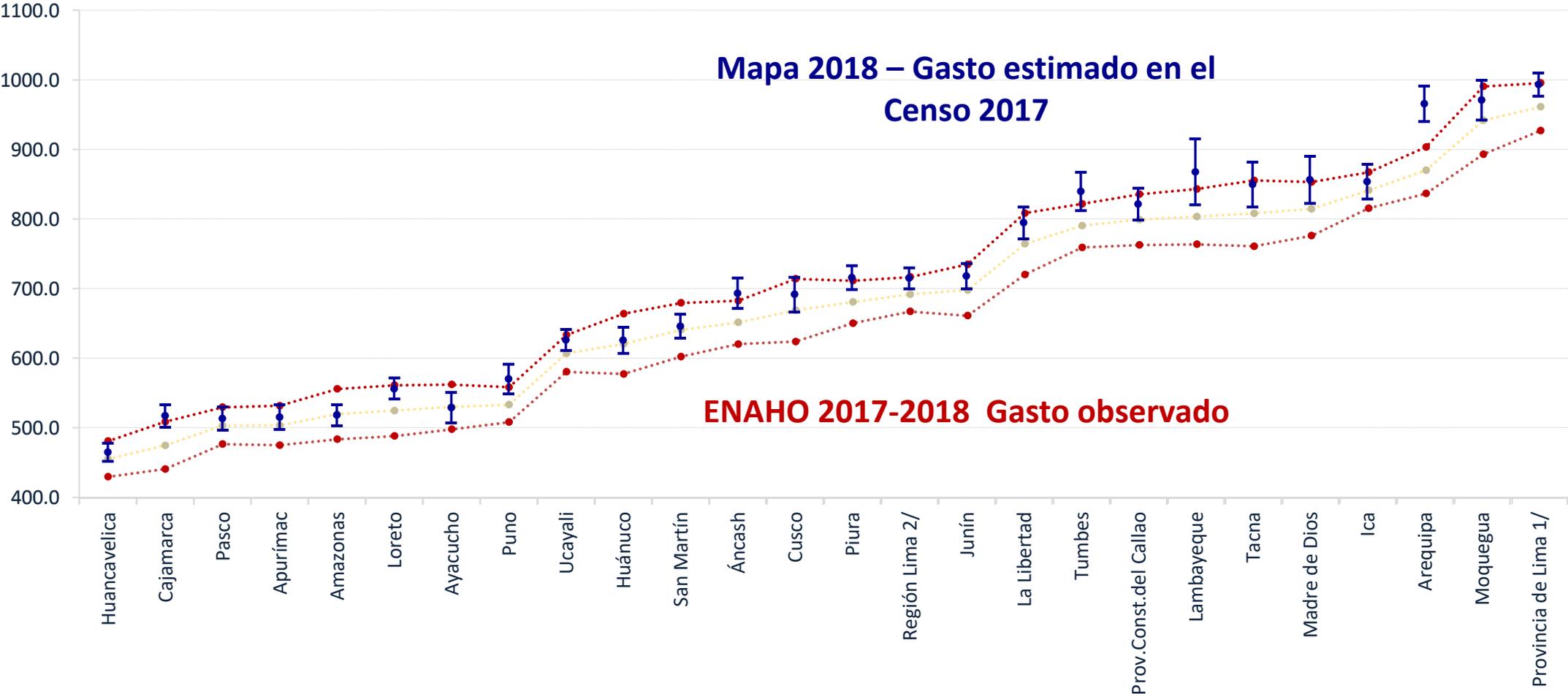


1/ Comprende los 43 distritos que conforman la provincia de Lima.
 2/ Comprende las Provincias de Barranca, Cajatambo, Canta, Cañete, Huaral, Huarochirí, Huaura, Oyón y Yauyos.
 Fuente: INEI – Mapa de Pobreza Provincial y Distrital 2018.

Validación interna – Sensibilidad en la muestra

GASTO OBSERVADO Y GASTO ESTIMADO DE LA MUESTRA

En el gráfico se presentan los gastos per cápita del hogar observados en la ENAHO 2017-2018 y los gastos per cápita del hogar estimados en el Censo 2017, donde todos los modelos se encuentran dentro del intervalo de confianza de la encuesta al 95%, a excepción de Arequipa, concluyendo que las estimaciones del gasto tienen buena precisión



1/ Comprende los 43 distritos que conforman la provincia de Lima.

2/ Comprende las Provincias de Barranca, Cajatambo, Canta, Cañete, Huaral, Huarochirí, Huaura, Oyón y Yauyos.

Fuente: INEI – Mapa de Pobreza Provincial y Distrital 2018.

- Se evaluó la capacidad predictiva de los modelos dividiendo la muestra de la ENAHO 2017-2018 en dos partes iguales.
- Se analizó la sensibilidad de los resultados de los modelos departamentales a distintas desagregaciones como área de residencia, dominio geográfico y región natural.

Validación externa consistió en la confrontación de los resultados de la pobreza monetaria con los registros administrativos a nivel distrital proporcionados por los gestores de los programas sociales.

5. Agrupación de distritos

Objetivo

Obtener una mayor precisión de incidencia de la pobreza monetaria total en aquellos distritos de menor tamaño y con alto coeficiente de variación (c.v.).

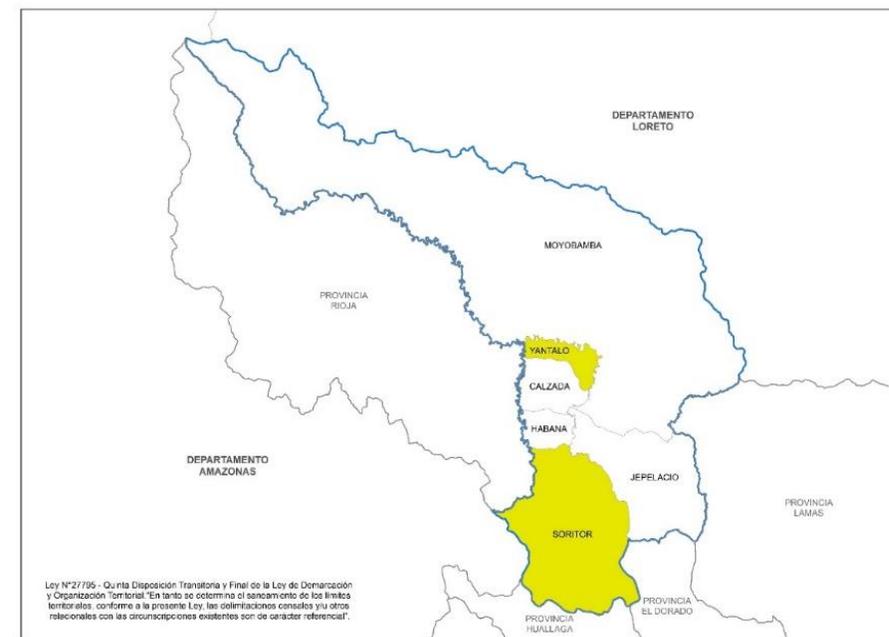
Para identificar los distritos a ser agrupados, se seleccionó aquellos con coeficiente de variación mayor al 25% y con menos de mil hogares.

En total se identificaron **107 distritos** para la agregación.

Criterios

- ✓ Cercanía del distrito
- ✓ Carácter urbano/rural del distrito y
- ✓ Pertenencia al mismo grupo robusto (18) dentro de la misma provincia.

Características	Grupo	
	61	61
Departamento	San Martín	San Martín
Provincia	Moyobamba	Moyobamba
Distrito	Soritor	Yantalo
Hogares	5 992	800
Área urbana	72.8	78.0
Área rural	27.2	22.0
Incidencia inicial		
Incidencia	33.0	25.8
C.V.	14.4	26.3
Incidencia agrupada		
Incidencia	31.1	31.1
C.V.	14.9	14.9

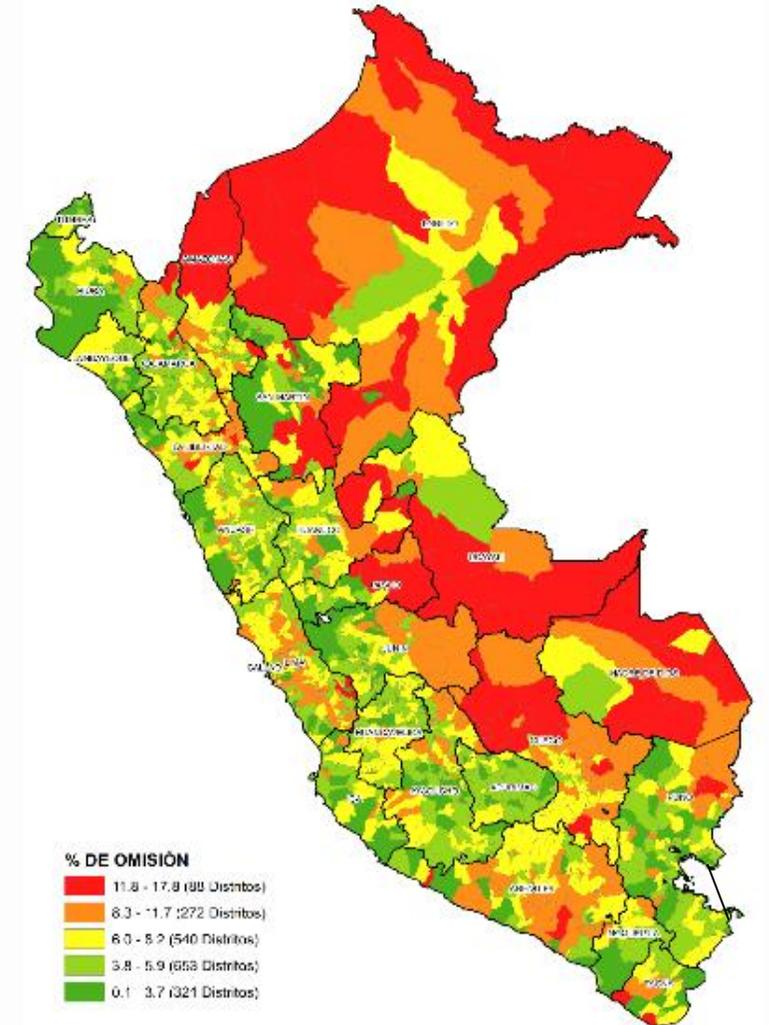


Limitaciones

El nivel de precisión de las estimaciones distritales de la incidencia de pobreza monetaria se ve afectado por el tamaño de la población total de los distritos. El 30% (570) de los distritos tienen una población total menor a 2 mil habitantes.

La no disponibilidad de otras fuentes de información relevantes para explicar el consumo del hogar, como registros administrativos de empleo, afecta la eficiencia de las estimaciones de pobreza.

La omisión de población en el censo no ha sido homogénea; se tienen distritos donde la omisión ha sido muy alta, sobre todo en la región Selva, donde 106 de ellos presentan una omisión superior al 10%, los cuales pueden generar sesgos en la estimación de la pobreza.

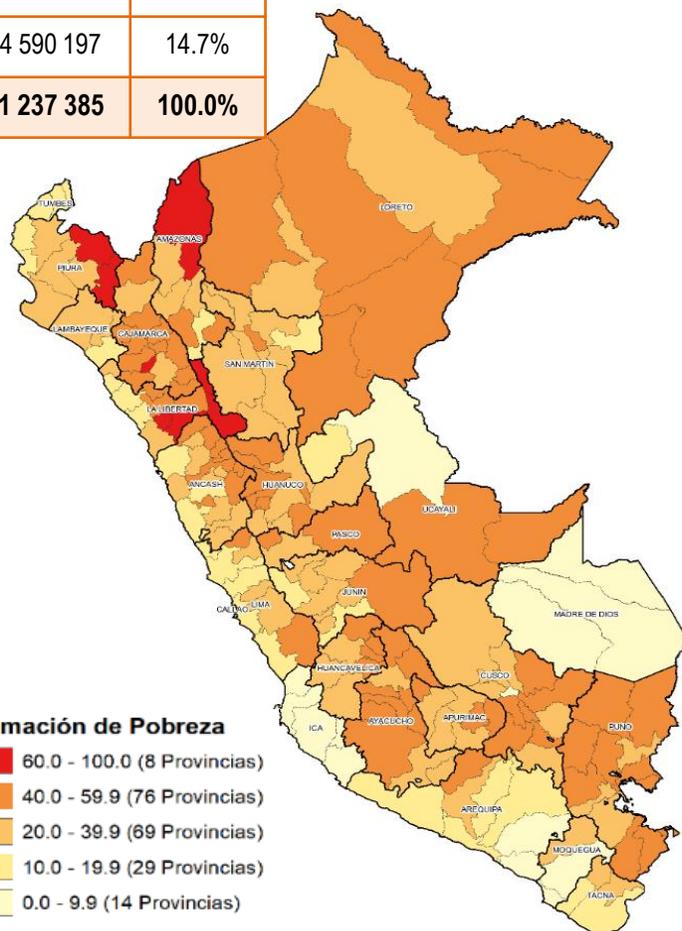


RESULTADOS

Estimación de la Pobreza total por provincias

Número de provincias, hogares y población por rango de pobreza total, 2018

Rango de pobreza total	Provincias		Hogares		Población	
	Cantidad	%	Cantidad	%	Cantidad	%
<20.0	43	21.9%	4 990 883	60.5%	19 135 054	61.3%
De 20 a 39.9	69	35.2%	1 977 737	24.0%	7 512 134	24.0%
>=40.0	84	42.9%	1 283 664	15.6%	4 590 197	14.7%
Total	196	100.0%	8 252 284	100.0%	31 237 385	100.0%



VEINTE PROVINCIAS MÁS POBRES, 2018

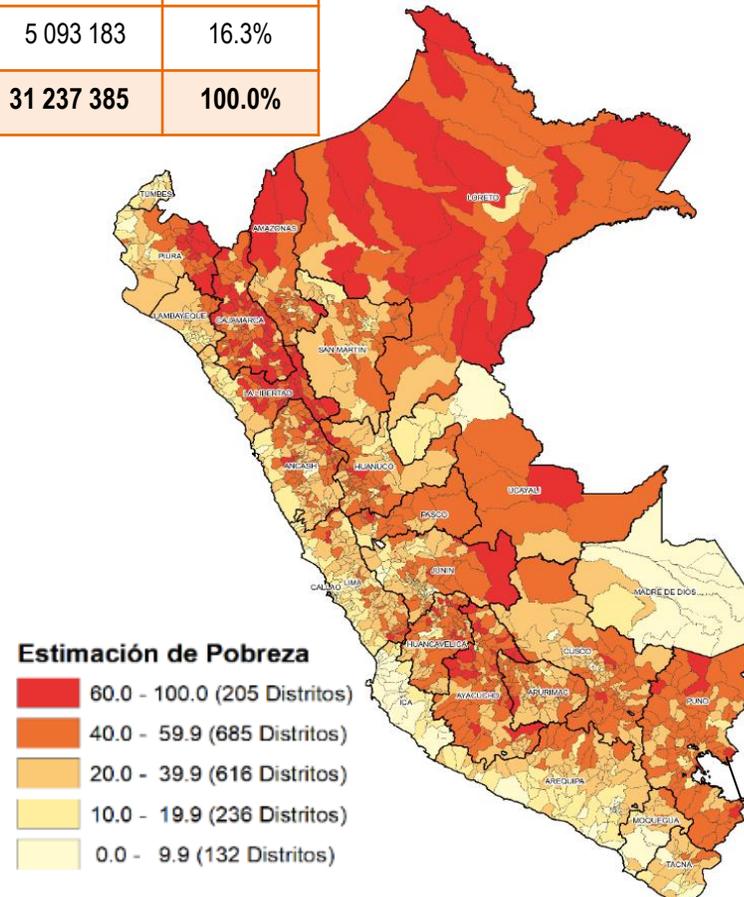
Ranking	Departamento	Provincia	Mapa 2018			
			Valor	I.C. al 95%		C.V.
				L.I.	L.S.	
1	Piura	Ayabaca	66.9	63.3	70.4	2.7
2	La Libertad	Bolívar	66.5	61.1	72.0	4.2
3	Piura	Huancabamba	65.0	61.2	68.9	3.0
4	Amazonas	Condorcanqui	64.4	59.8	69.0	3.7
5	La Libertad	Julcan	63.9	58.1	69.8	4.6
6	Cajamarca	San Pablo	63.6	56.9	70.4	5.4
7	La Libertad	Pataz	61.2	56.7	65.6	3.7
8	La Libertad	Santiago de Chuco	60.3	54.7	65.9	4.7
9	La Libertad	Sánchez Carrión	58.9	54.0	63.8	4.2
10	Pasco	Daniel Alcides Carrión	58.8	53.6	64.0	4.5
11	La Libertad	Otuzco	58.7	54.0	63.3	4.1
12	Huánuco	Huacaybamba	57.6	49.8	65.4	6.9
13	Loreto	Datem del Marañón	55.7	50.9	60.5	4.4
14	Cajamarca	Cutervo	55.3	51.1	59.5	3.8
15	Ayacucho	Víctor Fajardo	54.9	49.3	60.6	5.3
16	Áncash	Mariscal Luzuriaga	54.8	46.5	63.0	7.7
17	Ucayali	Purus	54.7	45.8	63.6	8.3
18	La Libertad	Gran Chimú	54.7	48.5	60.8	5.7
19	Cajamarca	San Miguel	54.7	50.2	59.2	4.2
20	Cajamarca	Celendin	54.2	50.8	57.6	3.2

* Provincia con estimación directa proveniente de la encuesta 2015-2018 válida.
Fuente: INEI – Mapa de Pobreza Provincial y Distrital 2018.

Estimación de la Pobreza total por distritos

Número de distritos, hogares y población por rango de pobreza total, 2018

Rango de pobreza total	Distritos		Hogares		Población	
	Cantidad	%	Cantidad	%	Cantidad	%
<20.0	368	19.6%	5 367 737	65.0%	20 671 270	66.3%
De 20 a 39.9	616	32.9%	1 476 144	17.9%	5 472 932	17.6%
>=40.0	890	47.5%	1 408 403	17.1%	5 093 183	16.3%
Total	1 874	100.0%	8 252 284	100.0%	31 237 385	100.0%



PERÚ: VEINTE DISTRITOS MÁS POBRES, 2018

N	Departamento	Provincia	Distrito	Mapa 2018			
				Valor	I.C. al 95%		C.V.
					L.I.	L.S.	
1	Ayacucho	Huanta	Uchuraccay	83.4	76.6	90.2	4.2
2	Cajamarca	Cutervo	Cujillo	79.8	67.0	92.7	8.2
3	Ayacucho	Huanta	Pucacolpa	79.8	70.3	89.3	6.1
4	Cajamarca	Celendin	Oxamarca	79.6	70.4	88.7	5.9
5	Piura	Ayabaca	Lagunas	79.5	72.4	86.5	4.5
6	Piura	Ayabaca	Sapillica	79.3	72.7	86.0	4.3
7	Cajamarca	Celendin	Huasmin	79.3	72.3	86.3	4.5
8	Cajamarca	Cajamarca	Cospan	79.0	67.0	91.0	7.7
9	La Libertad	Pataz	Huayo	78.6	70.8	86.4	5.0
10	La Libertad	Sanchez Carrión	Sanagoran	78.2	66.4	89.9	7.7
11	La Libertad	Sanchez Carrión	Chugay	77.4	68.7	86.1	5.7
12	Piura	Ayabaca	Pacaipampa	77.3	71.5	83.2	3.8
13	La Libertad	Sanchez Carrión	Marcabal	77.3	67.3	87.2	6.6
14	Cajamarca	Jaen	Sallique	76.9	66.5	87.2	6.9
15	La Libertad	Bolivar	Condormarca	76.8	67.6	86.1	6.1
16	Ayacucho	Huanta	Chaca	76.5	65.1	87.8	7.6
17	Ayacucho	La Mar	Oroncoy	76.3	65.0	87.7	7.6
18	La Libertad	Pataz	Chillia	76.3	68.0	84.6	5.6
19	Cajamarca	San Marcos	José Sabogal	76.0	67.5	84.5	5.7
20	Cajamarca	San Ignacio	Namballe	75.3	67.8	82.9	5.1

* Distrito con estimación directa proveniente de la encuesta 2015-2018 válida.
Fuente: INEI – Mapa de Pobreza Provincial y Distrital 2018.



www.inei.gob.pe



Instituto Nacional de
Estadística e Informática - Perú



@INEI_oficial



inei_peru



TV INEI Perú



inei_peru



INEI - Instituto Nacional de
Estadísticas e Informática