

Documento Metodológico

Mapa de Pobreza Monetaria Distrital¹

Paul A. Ervin²
Lyliana Gayoso²
Gustavo Enrique Galeano³
Celso Ovando³
Cristian Roa³
Eliana Rubiano-Matulevich²

2025
Asunción-Paraguay

¹ Los autores agradecen a Paul Corral, Jaime Estuardo Fernández y a Nancy Lozano por sus valiosas sugerencias.

² Banco Mundial

³ Instituto Nacional de Estadística de Paraguay

Tabla de Contenidos

Resumen ejecutivo	4
1. Introducción.....	6
2. Descripción de los datos	9
2.1 Delimitación administrativa de Paraguay	9
2.2 Encuesta Permanente de Hogares Continua (EPHC) 2023.....	10
2.2.1 Principales características	10
2.2.2 Cobertura geográfica de la encuesta	11
2.2.3 El agregado de bienestar y las líneas de pobreza.....	12
2.2.4 Los estimadores directos de la tasa de pobreza	13
2.3 El Censo Nacional de Población y Vivienda de 2022 (CNPV 2022)	14
2.4 Comparación de variables entre la EPHC 2023 y el CNPV 2022	15
3. Método de estimación.....	19
4. Desarrollo del modelo	21
4.1 Selección del nivel de los efectos por área	21
4.2 Transformación de la variable dependiente	24
4.3 Selección de un modelo nacional o varios modelos subnacionales	27
4.4 Selección de las variables a incluir en el modelo	28
4.5 Cumplimiento de los supuestos	29
4.6 Evaluación de las estimaciones.....	32
5. Resultados del modelo.....	34
5.1 Selección del modelo final.....	34
5.2 Métricas de ajuste y precisión.....	36
5.3 Comparación de los resultados con los estimadores directos	36
5.4 Comparación entre las áreas muestreadas en la encuesta y las no muestreadas	38
6. Mapa de pobreza.....	39
6.1 Mapa de pobreza utilizando el ingreso per cápita como medida de bienestar	39
6.2 Distribución territorial de la población en pobreza.....	41
7. Posibles extensiones	42
7.1 Implicaciones para la focalización de políticas públicas	43

7.2	Análisis de factores estructurales: pobreza monetaria y necesidades básicas insatisfechas	45
7.3	Otras posibles extensiones al mapa de pobreza	47
8.	Conclusiones.....	47
	Bibliografía	49
	Anexos	51
	Anexo A: Lista de variables candidatas y criterios de selección.....	51
	Anexo B: Estadísticas de cobertura de área	52
	Anexo C: Pruebas de normalidad para transformación logarítmica	53
	Estadísticas de asimetría y curtosis	53
	Prueba de asimetría y curtosis	53
	Prueba W de Shapiro-Wilk	53
	Gráfico cuantil-cuantil con transformación logarítmica del ingreso	53
	Anexo D: Pruebas de normalidad para transformación Box-Cox.....	54
	Estadísticas de asimetría y curtosis	54
	Prueba de asimetría y curtosis	54
	Prueba W de Shapiro-Wilk	54
	Gráfico cuantil-cuantil con transformación Box-Cox del ingreso.....	54
	Anexo E: Pruebas de normalidad para transformación logarítmica de sesgo cero	55
	Estadísticas de asimetría y curtosis	55
	Prueba de asimetría y curtosis	55
	Prueba W de Shapiro-Wilk	55
	Gráfico cuantil-cuantil con transformación logarítmica de sesgo cero del ingreso	55
	Anexo F: ECM de modelos alternativos	56
	ECM Modelo Nacional.....	56
	ECM Modelos Departamentales.....	57
	Anexo G: Lista de variables y sus definiciones	60
	Anexo H: Diagramas de dispersión.....	62
	Anexo I: Comparación de ratios a nivel departamental	63

Resumen ejecutivo

Este reporte presenta el mapa de pobreza monetaria de Paraguay para 2022, utilizando metodologías de Estimación en Áreas Pequeñas para generar estimaciones de pobreza a nivel distrital para los 17 departamentos y a nivel barrial para el distrito capital del país, Asunción. El estudio combina datos de la Encuesta Permanente de Hogares Continua 2023 con el Censo Nacional de Población y Vivienda 2022, empleando el Mejor Estimador Empírico del Censo para mantener consistencia con las mediciones oficiales mientras genera estimaciones precisas para 263 distritos del país y los 67 barrios poblados de Asunción.

Este análisis se enfoca exclusivamente en la población no indígena de Paraguay. La población indígena, (2,3% del total) fue censada mediante un operativo específico con cuestionario y variables diferenciadas, diseñadas para capturar sus formas de organización social, económica y territorial. Estas variables no son comparables con las del censo nacional, y la Encuesta Permanente de Hogares Continua (EPHC) tampoco incluye población indígena. Por tanto, el presente mapa de pobreza no cubre a la población indígena, cuyo análisis requiere marcos analíticos propios que deberán desarrollarse en futuros trabajos.

Los resultados revelan amplias disparidades territoriales que trascienden los promedios departamentales tradicionalmente reportados. Al 2022, el distrito más pobre del país es Paso Barreto, en el departamento de Concepción, con una tasa de pobreza del 53,3%, mientras que el distrito menos pobre es Fernando de la Mora, con una tasa de pobreza de 3,7%. Esta diferencia de más de 49 puntos porcentuales refleja la heterogeneidad socioeconómica que caracteriza al territorio paraguayo y evidencia la necesidad de políticas públicas diferenciadas territorialmente.

Los patrones geográficos revelan la influencia de factores estructurales en la distribución de la pobreza, confirmada mediante el análisis comparativo con indicadores de necesidades básicas insatisfechas. En la Región Occidental, Alto Paraguay registra la tasa de pobreza departamental más elevada del país (38,7%) y una convergencia crítica entre pobreza monetaria y carencias estructurales. Cabe señalar que las estimaciones para los departamentos de Alto Paraguay y Boquerón son sintéticas, ya que estas áreas no fueron incluidas en la muestra de la EPHC 2023. Por tanto, sus valores se derivan exclusivamente del modelo aplicado sobre variables censales, sin información directa de encuesta, y por ello sus resultados deben

interpretarse con cautela. En contraste, en la Región Oriental, el departamento Central (14,7%) y el distrito de Asunción (7,3%) muestran niveles considerablemente menores niveles de pobreza, aunque con patrones heterogéneos que evidencian diferentes combinaciones entre limitaciones de ingresos y deficiencias en infraestructura básica. Incluso Asunción presenta alta variación interna, con tasas que oscilan entre 0,03% en el barrio Mburicao y 42,5% en Banco San Miguel, evidenciando procesos de segregación socioespacial urbana.

El desafío aún es considerable: aproximadamente 1,3 millones de paraguayos viven en condiciones de pobreza monetaria, representando el 22,7% de la población nacional no indígena en 2023. El análisis territorial revela que, si bien las tasas de pobreza son menores en centros urbanos, la concentración absoluta de población pobre se localiza principalmente en estos territorios, planteando desafíos diferenciados para el diseño de políticas públicas que deben abordar tanto la intensidad del problema en áreas rurales de baja densidad como la magnitud absoluta en centros urbanos y periurbanos.

Los hallazgos presentados en este documento proporcionan evidencia cuantitativa robusta para orientar la focalización territorial de programas sociales, la priorización de inversiones en infraestructura y servicios básicos, y el diseño de intervenciones adaptadas a las realidades locales específicas. El desarrollo de tipologías territoriales basadas en la combinación de incidencia de pobreza y concentración poblacional facilita la identificación de estrategias diferenciadas, desde programas integrales para áreas de máxima prioridad hasta intervenciones específicas según el contexto territorial. Los mapas de pobreza constituyen herramientas fundamentales para superar las limitaciones de enfoques basados únicamente en promedios departamentales y avanzar hacia políticas de desarrollo territorial más efectivas y equitativas.

1. Introducción

La recopilación de datos de bienestar que permitan estimar tasas de pobreza y desigualdad representa un desafío considerable, particularmente en países en desarrollo, debido a los altos costos, largos tiempos de implementación y complejos requerimientos técnicos (Lanjouw, 2021; Yoshida, et al., 2021). Como respuesta, la mayoría de países de ingresos medios y bajos implementan encuestas de hogares como fuente principal para generar indicadores de condiciones de vida. Sin embargo, estas encuestas típicamente proporcionan representatividad únicamente a nivel nacional y para algunas divisiones administrativas mayores, como áreas de residencia (urbano/rural), regiones o departamentos.

Para tener una mejor comprensión de las disparidades económicas y sociales que existen a nivel territorial y que a niveles administrativos más agregados no se pueden visualizar, idealmente se deberían contar con datos representativos al nivel administrativo más bajo posible, como distritos, barrios, etc. Contar con datos de pobreza y desigualdad más desagregados permite diseñar intervenciones mejor adaptadas a las realidades locales. En este sentido, si bien los censos representan la fuente de información más completa, ya que contienen información de todos los hogares y las personas del país, típicamente éstos no recogen información sobre medidas de bienestar, como ingreso o gastos, lo que representa una limitación para la construcción de medidas de pobreza y desigualdad en base a datos censales.

Estas limitaciones metodológicas son comunes en la región y han motivado importantes avances técnicos en los últimos años. Se han desarrollado diferentes metodologías de imputación que permiten aprovechar la información complementaria de encuestas y censos para la estimación de tasas de pobreza a nivel local. En particular, estas limitaciones han dado lugar a una amplia literatura sobre estimación para áreas pequeñas (SAE, por sus siglas en inglés) para el mapeo de la pobreza. La técnica SAE combina información auxiliar de censos, registros u otras encuestas más amplias para generar estadísticas desagregadas a nivel de áreas pequeñas con suficiente precisión cuando los datos de la encuesta por sí solos resultan insuficientes (Corral et al., 2022). Esta técnica consiste en utilizar un conjunto de datos, la encuesta de hogares, para entrenar un modelo que luego es utilizado para imputar ingresos (o gastos) en otro conjunto de

datos, por ejemplo, el censo. Esta metodología ganó prominencia en las últimas décadas, en particular con el modelo lineal propuesto por Fay y Herriot (1979), y el modelo de error anidado propuesto por Battese et al. (1988). La literatura en esta área avanzó con la innovación introducida por Elbers, Lanjouw, y Lanjouw (2003), los que extendieron el modelo propuesto por Battese et al. (1988), permitiendo la imputación del bienestar basada en el uso de microdatos, lo cual a su vez permite el cálculo de numerosos indicadores de pobreza y bienestar más allá de los agregados de área, y aplicando el modelo Alfa para corregir por heterocedasticidad. De esta forma, el modelo conocido como ELL, se convirtió en referencia en la elaboración de mapas de pobreza. Luego, Molina y Rao (2010), introdujeron el mejor estimador empírico del censo (Census Empirical Best, o Census EB, por sus siglas en inglés), el cual se caracteriza por utilizar más eficientemente los datos de encuestas disponibles y por abordar las preocupaciones sobre la precisión de las estimaciones de pobreza (Banerjee et al., 2006). Este enfoque mejora la metodología de Elbers, Lanjouw y Lanjouw (2003), ampliamente utilizada antes del 2015. Más recientemente, Corral et al. (2021) actualizaron el método al adaptar el procedimiento de simulación de Monte Carlo de Molina y Rao (2010) y la metodología de Van der Weide (2014). Esto les permitió abordar la heteroscedasticidad generada por los efectos de área e incorporar los ponderadores de la encuesta.

En Paraguay, las tasas oficiales de pobreza monetaria se calculan anualmente desde 1997/1998 utilizando como fuente primaria la Encuesta Permanente de Hogares (EPH), posteriormente reemplazada por la Encuesta Permanente de Hogares Continua (EPHC) en 2017. El agregado de bienestar adoptado por el país es el ingreso per cápita mensual del hogar. A pesar de la disponibilidad de datos de forma anual de la EPHC, la misma presenta limitaciones. Si bien, la EPHC tiene representatividad nacional, por área de residencia y para departamentos selectos, la misma no permite inferencias estadísticamente válidas para niveles administrativos menores. Por otra parte, el Censo Nacional de Población y Vivienda 2022 (CNPV 2022) recopiló información detallada de todos los hogares e individuos del país, pero no incluye datos sobre ingresos que permitan estimar directamente la pobreza y desigualdad. Esta configuración de datos presenta una oportunidad metodológica ideal para aplicar técnicas de estimación en áreas pequeñas, aprovechando la cobertura exhaustiva del censo y el detalle de bienestar de la encuesta.

En este reporte, presentamos las estimaciones de tasas de pobreza monetaria en Paraguay al nivel administrativo más bajo posible: distritos del país y barrios del distrito capital, Asunción, complementadas con un análisis territorial integral de los patrones de distribución espacial. Las tasas de pobreza a nivel distrital existentes en el país datan del 2002, y datos recientes de la EPHC denotan un cambio en las tasas de pobreza monetaria a nivel departamental, reflejando un posible cambio en niveles más desagregados. Además, desde el 2002, la estructura administrativa a nivel distrital ha cambiado en Paraguay, pasando de 236 distritos (Robles y Santander, 2004) a 263 distritos actualmente, incluyendo el distrito capital del país. Por todo esto, contar con estimaciones más recientes y con análisis territoriales diferenciados es de gran relevancia para informar a los hacedores de política de las nuevas realidades locales y de la necesidad de diseñar e implementar intervenciones que busquen reducir la pobreza ajustadas a estas realidades específicas.

Para la estimación de los mapas de pobreza en este estudio, se combinan los datos del CNPV 2022 y de la EPHC 2023. Se optó por el uso de la EPHC 2023 debido a su proximidad temporal al censo, y por la minimización de desajustes que esta base permite en términos de indicadores de bienestar. Esta proximidad temporal permitió aplicar la metodología SAE y el mejor estimador empírico del censo o Census EB. Cabe señalar que, este análisis se enfoca exclusivamente en la población no indígena de Paraguay. La población indígena, que representa el 2,3% del total, fue censada mediante un operativo específico con cuestionario y variables diferenciadas, diseñadas para capturar sus formas de organización social, económica y territorial. Estas variables no son comparables con las del censo nacional y, además, la EPHC no incluye población indígena. Por lo tanto, el presente mapa de pobreza no cubre a la población indígena, cuyo análisis requiere marcos analíticos propios que deberán desarrollarse en futuros trabajos.

El análisis presentado en este estudio avanza esfuerzos previos de SAE de dos maneras importantes. Primero, utiliza el censo más reciente de Paraguay CNPV 2022 y la EPHC 2023 para generar estimaciones de la distribución espacial de la pobreza para la división administrativa actualizada de Paraguay de 263 distritos del país y 67 barrios poblados del distrito capital del país, Asunción. Segundo, incorpora mejoras metodológicas mediante el empleo del marco de predictores empíricos óptimos (EB) propuesto por Molina y Rao (2010). El análisis

sigue las Directrices para la Estimación de Áreas Pequeñas para el Mapeo de la Pobreza desarrolladas por Corral, Molina, Cojocarú y Segovia (2022), y presenta un conjunto de indicadores de pobreza monetaria, incluyendo la tasa de recuento (FGT0), el índice de brecha de pobreza (FGT1), el índice de severidad de la pobreza (FGT2) y el número estimado de individuos pobres. Además, el estudio complementa las estimaciones de pobreza con análisis territoriales que facilitan la identificación de tipologías para la focalización de políticas públicas y exploran la relación entre pobreza monetaria y factores estructurales.

El resto del documento se organiza de la siguiente manera: La Sección 2 presenta la descripción de las bases de datos utilizadas en el análisis, seguida por la Sección 3 que describe la metodología de estimación empleada. La Sección 4 explica la aplicación de dicho método para el desarrollo del modelo, mientras que la Sección 5 presenta los resultados del modelo final. Por su parte, la sección 6 presenta los mapas de ingreso y pobreza, y la sección 7 describe posibles extensiones de los mapas. Finalmente, la Sección 8 resume los resultados.

2. Descripción de los datos

La aplicación del método de Estimación en Áreas Pequeñas (SAE) requiere la disponibilidad de dos fuentes de datos complementarias: una encuesta de hogares que contenga información detallada sobre bienestar y un censo poblacional con cobertura universal. En el caso de Paraguay, esta condición se cumple mediante la combinación de la Encuesta Permanente de Hogares Continua (EPHC) 2023 y el Censo Nacional de Población y Vivienda 2022. La proximidad temporal entre ambas fuentes, junto con la compatibilidad de sus variables, permite la implementación de metodologías SAE con estándares técnicos apropiados. Esta sección describe las características principales de ambas fuentes de datos, analiza su compatibilidad y presenta los criterios empleados para la selección de variables del modelo.

2.1 Delimitación administrativa de Paraguay

El Paraguay está dividido en dos regiones, las cuales están separadas por el Río Paraguay, con diferencias topográficas marcadas. La Región Occidental, situada al norte del río Paraguay y conocida también como Chaco o Chaco Boreal, se caracteriza por una extensa llanura con abundante vegetación y una fauna variada. Esta región abarca 3 departamentos de gran superficie

y alberga aproximadamente el 3% de la población total. La Región Oriental, situada al sur y comprendida entre los ríos Paraguay y Paraná, se caracteriza por llanuras fértiles, ríos, arroyos, lagos, y amplios bosques con praderas (Ministerio de Asuntos Exteriores, Unión Europea y Cooperación, 2024). Esta región abarca 14 departamentos y el distrito capital Asunción, concentrando al 97% de la población total. El país cuenta con 17 departamentos que se subdividen en distritos. En el 2022, existían un total de 263 distritos, incluyendo la capital Asunción, que se subdivide en 68 barrios, de los cuales 67 están poblados.

2.2 Encuesta Permanente de Hogares Continua (EPHC) 2023

Para la estimación de pobreza en áreas pequeñas se requiere idealmente una encuesta y un censo contemporáneos, implementados en el mismo periodo o lo más cercanos posible temporalmente. En el caso de Paraguay, el último censo data del 2022, y si bien se cuenta con la EPHC del mismo año, se determinó utilizar la EPHC 2023 por las siguientes ventajas comparativas:

1. Cercanía temporal: La EPHC 2023 refleja los ingresos de los hogares del año agrícola 2022, coincidente con el período de referencia del censo.
2. Minimización de desajustes: La cercanía temporal minimiza posibles desajustes en la medición de indicadores de bienestar.
3. Validez de resultados: La correspondencia temporal fortalece la validez de los resultados obtenidos a través de la metodología de estimación en áreas pequeñas (SAE).

La EPHC 2023 y sus predecesoras son encuestas de hogares representativas a nivel nacional diseñadas para recopilar datos detallados sobre las condiciones de vida y los ingresos de los hogares. Las mismas sirven como fuente primaria para la construcción del agregado de bienestar, el ingreso per cápita mensual del hogar, y para la estimación de las tasas de pobreza oficiales, proporcionando información valiosa sobre el nivel de vida en todo el país.

2.2.1 Principales características

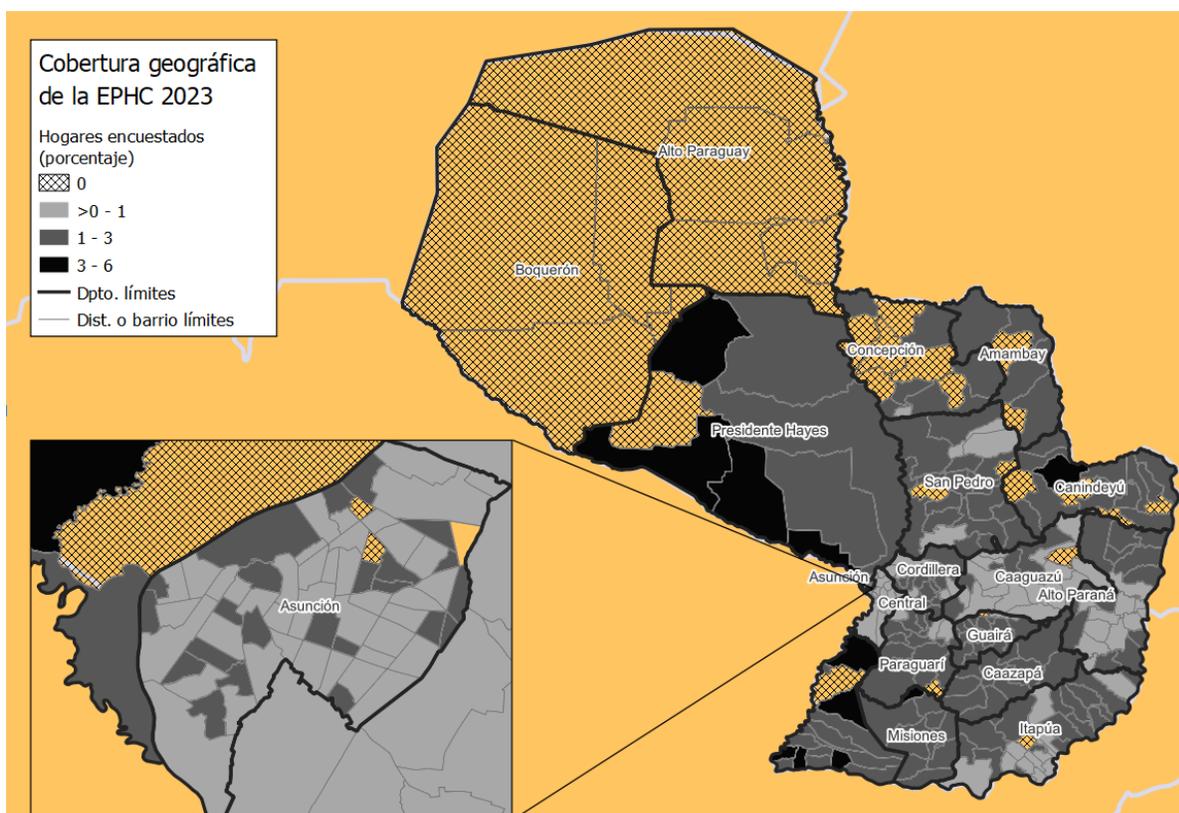
La EPHC 2023 anual está diseñada para ser representativa además del nivel nacional, por áreas de residencia (urbano/rural), y a nivel de 15 departamentos del país (ver más abajo cobertura geográfica). La misma se llevó a cabo entre enero y diciembre del 2023, y contiene un

total de 58.005 personas y 17.037 hogares, representando a un total de 5.885.678 personas y 1.666.285 hogares particulares con personas presentes según el Censo 2022, respectivamente.

2.2.2 Cobertura geográfica de la encuesta

La EPHC 2023 tiene una cobertura geográfica de 65 de los 67 barrios poblados de la capital de Asunción⁴, y de 234 distritos, lo que representa un 89% del total de distritos del país. Todos estos distritos se encuentran ubicados en 15 departamentos del país. La EPHC 2023 no se aplicó en los distritos de los departamentos de Alto Paraguay y Boquerón, los que se sitúan en la Región Occidental del país. La Figura 2-1 a continuación, refleja la cobertura geográfica de la EPHC 2023, denotando además el porcentaje de hogares encuestados por distrito, y en el caso de la capital de Asunción, el porcentaje de hogares por barrios.

Figura 2-1 Mapa de cobertura geográfica de la EPHC 2023



Fuente: Elaboración propia en base a datos de la EPHC 2023.

Notas: La EPHC no incluye los departamentos de Boquerón y Alto Paraguay, ni las comunidades indígenas ni las viviendas colectivas. En consecuencia, las estimaciones para Alto Paraguay y Boquerón presentadas en este documento son sintéticas y deben interpretarse con cautela, al derivarse únicamente de las variables censales sin información directa de la encuesta.

⁴ El barrio Ñu Guasú de la ciudad de Asunción es un barrio localizado en un parque nacional, por lo que no está habitado.

2.2.3 El agregado de bienestar y las líneas de pobreza

El agregado de bienestar en Paraguay es el ingreso per cápita mensual del hogar. La EPHC 2023 recoge información sobre las distintas fuentes de ingreso de los hogares, representados en moneda local, guaraníes. A partir de esta información, el equipo técnico del INE realiza imputaciones en casos necesarios, y un ajuste espacial a nivel de área de residencia, como también ajustes temporales. Luego de las imputaciones y ajustes, se calcula el ingreso mensual total del hogar, para finalmente computar el ingreso mensual per cápita de los hogares.

Por su parte, las líneas de pobreza son valores monetarios, representados en la moneda local, guaraníes. El INE estima dos líneas de pobreza monetaria: la línea de pobreza total (LPT) y la línea de pobreza extrema (LPE). La LPT refleja el costo monetario mensual de una canasta básica de alimentos y otros bienes y servicios esenciales, en tanto que la LPE refleja el costo monetario mensual de una canasta básica de alimentos (INE, 2024). El cálculo de las líneas de pobreza se realiza a partir de dos fuentes. Por un lado, se utiliza la EPHC para obtener los ingresos de los hogares, y por otro, se utiliza la estructura de consumo de la Encuesta de Ingresos y Gastos 2011/2012 para obtener el costo de la canasta básica de alimentos, cuyos valores son actualizados anualmente con el Índice de Precios al Consumidor (IPC). Es importante destacar que las líneas de pobreza difieren para las áreas urbanas y rurales, ya que las mismas reflejan las diferencias en el costo de vida que se experimentan en ambas áreas (INE, 2024).

A continuación, se presentan las líneas de pobreza oficiales para el 2023:

Tabla 1 – Líneas de pobreza por área geográfica en guaraníes

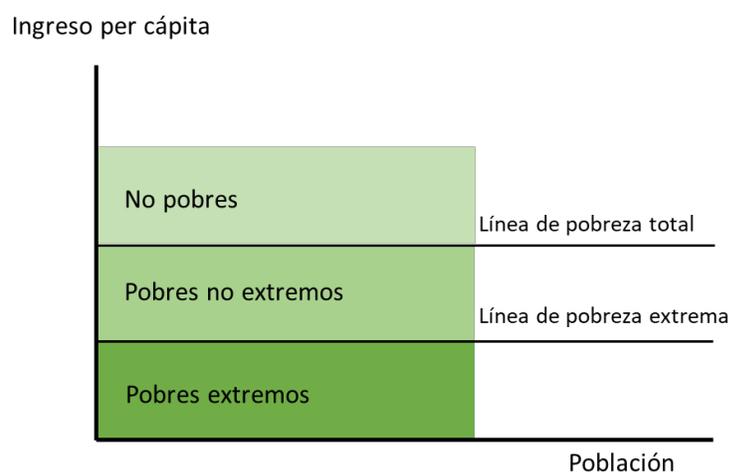
Urbana		Rural	
Línea de Pobreza Extrema	Línea de Pobreza Total	Línea de Pobreza Extrema	Línea de Pobreza Total
359.480	853.778	328.289	618.789

Fuente: INE (2024)

2.2.4 Los estimadores directos de la tasa de pobreza

Para el cálculo y estimación de las tasas de pobreza, total y extrema, primero se identifica la condición de pobreza de los hogares a través de la comparación entre el ingreso per cápita mensual del hogar y las líneas de pobreza, de manera a que un hogar se encuentra en situación de pobreza si su ingreso per cápita mensual es menor a la línea de pobreza. El siguiente gráfico refleja este procedimiento:

Figura 2-2 – Aplicación de líneas de pobreza



Fuente: INE (2024)

A partir de la identificación de la condición de pobreza del hogar, se calcula el porcentaje de personas que se encuentra en situación de pobreza para estimar las tasas de pobreza a nivel nacional, departamental, y por áreas, como se presenta en la tabla a continuación:

Tabla 2 – Estimadores directos de pobreza en Paraguay 2023

Nivel	Pobreza Total (%)	Intervalo de Confianza del 95%		Población	Hogares
		Límite Inferior	Límite Superior		
Departamento					
Asunción	9,7%	7,1%	13,1%	432.995	124.227
Concepción	35,9%	30,0%	42,2%	210.720	57.299
San Pedro	32,4%	28,4%	36,8%	359.032	105.601
Cordillera	23,9%	19,2%	29,4%	263.463	75.168
Guairá	34,0%	27,9%	40,7%	199.326	59.518
Caaguazú	31,9%	26,8%	37,4%	445.197	128.908
Caazapá	36,8%	28,8%	45,6%	156.364	46.641
Itapúa	29,8%	25,9%	34,0%	439.538	133.837
Misiones	26,4%	22,2%	31,2%	108.279	33.045
Paraguarí	30,5%	25,4%	36,0%	196.837	56.616
Alto Paraná	23,3%	20,3%	26,6%	724.225	206.962
Central	13,5%	11,7%	15,5%	1.858.846	495.230
Ñeembucú	26,0%	20,3%	32,7%	70.841	24.884
Amambay	22,4%	18,7%	26,5%	148.076	43.910
Canindeyú	35,1%	28,8%	42,1%	170.558	50.201
Presidente Hayes	18,8%	15,6%	22,5%	87.467	24.186
Boquerón	No conforman parte de la muestra de la EPHC				
Alto Paraguay					
Área de residencia					
Urbana	19,0%	17,6%	20,3%	3.649.916	1.026.175
Rural	28,8%	27,0%	30,6%	2.221.848	640.078
Total país					
Nacional	22,7%	21,6%	23,7%	5.871.764	1.666.233

Fuente: Elaboración propia con datos de la EPHC 2023 anual.

Notas: La EPHC no incluye los departamentos de Boquerón y Alto Paraguay, ni las comunidades indígenas ni las viviendas colectivas. Por tanto, las estimaciones para Alto Paraguay y Boquerón son sintéticas y deben interpretarse con cautela.

2.3 El Censo Nacional de Población y Vivienda de 2022 (CNPV 2022)

El CNPV 2022 es el octavo censo realizado en Paraguay, precedido por los censos de 1950, 1962, 1972, 1982, 1992, 2002, y 2012. El mismo tuvo lugar en noviembre del 2022 mediante dos operativos censales: el censo nacional para la población general no indígena y el censo indígena, cada uno con metodologías y cuestionarios específicos para sus respectivos universos poblacionales. La metodología aplicada en el trabajo de campo fue el “Hecho o Facto”, lo que implica que se registran los datos de todas las personas que pasaron la hora 0 (cero) en la vivienda el día del censo. Es decir, se considera a todas las personas que durmieron en la vivienda censada la noche anterior al día del censo. Por otra parte, los datos en áreas urbanas fueron recolectados durante un día, en tanto que en áreas rurales el periodo de recolección de

datos fue extendido hasta dos semanas. El formato del cuestionario aplicado fue en forma de papel, el cual estuvo estructurado en nueve capítulos, de acuerdo al siguiente esquema:

Tabla 3 – Estructura del cuestionario censal

Capítulo	Descripción
A	Localización geográfica
B	Datos de la Vivienda
C	Datos del Hogar y Mortalidad
D	Listado de Personas que Pasaron la Noche en el Hogar
E	Datos de la Población
F	Idioma y Educación
G	Migración
H	Estado Civil y Actividad Laboral
I	Fecundidad

De acuerdo con los resultados del Censo Nacional de Población y Viviendas 2022, 6.066.227 personas corresponden a 1.770.755 hogares particulares, incluyendo la población indígena, como se muestra en la tabla a continuación:

Tabla 4 – Número de hogares y población resultante del CNPV 2022

Hogares		Población	
Hogares Particulares con Personas Presentes (Nacional + Indígena)	Hogares Particulares con Personas Presentes Censo Nacional	Población en Hogares Particulares con Personas Presentes (Nacional + Indígena)	Población en Hogares Particulares con Personas Presentes Censo Nacional
1.770.755	1.733.659	6.066.227	5.928.804

Para la elaboración del presente mapa de pobreza se utilizó únicamente la información del censo nacional, excluyendo el operativo censal indígena, dado que sus variables no son comparables con las del cuestionario general.

2.4 Comparación de variables entre la EPHC 2023 y el CNPV 2022

Considerando que las estimaciones de áreas pequeñas se basan en datos a nivel de hogar para obtener la distribución del ingreso en función de características observables, es importante identificar las variables que sean comunes entre el censo y la encuesta, y que potencialmente expliquen el ingreso, y, por ende, la condición de pobreza, tanto a nivel de individuo como de hogar. Estas variables deben ser similares tanto semántica como analíticamente, es decir, las características de las variables en ambos conjuntos de datos (censo y encuesta) deben ser similares y presentar distribuciones similares. De lo contrario, las mismas podrían estar sesgadas y ser poco útiles (Corral, et al., 2022).

En este estudio, se compararon cuidadosamente cada pregunta del censo y de la EPHC 2023, de manera a evaluar su similitud semántica para luego realizar una comparación en base a las características distribucionales. En este proceso, se identificaron 33 preguntas del censo que presentaban coincidencias con la EPHC. Para identificar las variables que se incluirían como posibles candidatas para la estimación de pobreza en áreas pequeñas, se tuvieron en cuenta tres criterios:

1. Las preguntas y las opciones de respuestas deben ser las mismas o poder reclasificarse de manera similar para lograr una coincidencia directa.
2. Existe suficiente variación en las variables para explicar los niveles de ingresos. Por ejemplo, se omitieron las variables en las que el 5% o menos de la muestra no presenta esa característica.
3. Las medias de las variables son generalmente similares.

Con base al criterio (1), se eliminaron 6 variables ya que las mismas no reflejaban el mismo concepto. Estas incluían: i) la conexión a internet (esta variable se refería a la conexión domiciliaria en la EPHC, mientras que en el censo se refería a cualquier tipo de conexión a internet, incluyendo celular); (ii) posesión de TV (el censo hacía referencia a pantalla plana mientras que en la EPHC la pregunta se refería a cualquier tipo de TV); (iii) la tenencia de lavarropas (el censo hacía referencia a lavarropas automática mientras que la EPHC considera cualquier tipo de lavarropas); (iv) la tenencia de termocalefón (en el caso del censo la distinción entre termocalefón y ducha eléctrica no fue adecuadamente establecida, lo que derivó en registros que indicaban la presencia de termocalefón cuando, en realidad, correspondían a duchas eléctricas, mientras que en la EPHC solo se hace referencia al termocalefón, y no a su finalidad); (v) el idioma hablado (en el censo se pregunta si el miembro del hogar puede hablar cualquiera de los idiomas con respuestas múltiples permitidas, mientras que en la EPHC se hace referencia al idioma principal); y (vi) el registro civil (no pudo coincidir debido a la codificación diferenciada en el censo).

Según el criterio (2), se descartaron dos variables: el tipo de vivienda (96% respondió "casa, rancho") y tenencia de celular (95% respondió tener uno), debido a la baja variabilidad en las respuestas.

Para el criterio final (3), primero se examinó la relación entre las variables categóricas y el ingreso per cápita mensual con el objetivo de crear categorías que predigan la separación entre los niveles de ingreso y evitar el sobreajuste en muchas categorías que podrían cubrir un segmento pequeño de la población. Por ejemplo, los hogares más ricos generalmente reportaron tener pisos de bloque de cemento o ladrillo en comparación con otras categorías, por lo que estas categorías se agruparon para crear un indicador único sobre este tipo de material. Esta misma lógica se siguió para la creación de otras variables categóricas, como techo, pared, agua, drenaje, combustible para cocinar, recolección de basura, tipo de terreno de la casa, estructura del hogar (es decir, unipersonal, compuesto, etc.), estado civil y tipo de empleo del jefe de hogar (empleado remunerado, no remunerado, desempleado, inactivo jubilado o pensionado, u otros inactivos), entre otros. En este proceso también se crearon variables para capturar la demografía del hogar, como el número de mujeres en el hogar, el número de niños menores de 10 años, el número de jóvenes de 10 a 24 años, el número de adultos de 25 a 65 años y el número de personas mayores de 65 años y más.

La Tabla 5 a continuación presenta las medias nacionales en el censo y la EPHC para las variables consideradas en el modelo SAE. Con base en este análisis, además, se calculó la ratio de las medias de la EPHC versus del Censo para cada variable candidata. Finalmente, se optó por excluir las variables del censo cuyas ratios con respecto a las medias de la encuesta estaban fuera del rango 0,75- 1,25.⁵ Esto incluía el tipo de piso, tipo de desagüe, y si el hogar tiene una tableta. Por otra parte, se exploró además las diferencias de medias estandarizadas, es decir, la *d* de Cohen, la cual mide la magnitud de las diferencias entre dos medias en términos de desviaciones estándar. Todas las variables restantes como variables candidatas presentan magnitudes bajas de diferencias estandarizadas (por ej., inferiores a 0,2 desviaciones estándar).

⁵ Este rango fue establecido en base a que la mayoría de las ratios se encontraban dentro de este rango y las diferencias en las medias estandarizadas para las variables en este rango eran pequeñas (menores o muy cercanas a 0,2).

Tabla 5 – Comparación de medias entre variables de la EPHC 2023 y Censo 2022

Variables	EPHC 2023			Censo 2022	Métricas	
	Media	Intervalo de Confianza del 95%		Media	Ratio Censo / EPHC	Diferencias de medias estandarizadas
		Límite Superior	Límite Inferior			
g_tipo_pared1	0,851	0,841	0,861	0,838	0,984	0,04
g_tipo_piso1	0,647	0,635	0,659	0,537	0,830	0,22
g_tipo_techo1	0,552	0,538	0,565	0,534	0,968	0,04
g_agua_utilizada1	0,772	0,758	0,786	0,847	1,098	0,21
g_agua_medio1	0,907	0,900	0,913	0,874	0,964	0,10
g_tipo_desague1	0,098	0,089	0,109	0,127	1,295	0,09
g_tipo_combustible1	0,782	0,773	0,791	0,730	0,933	0,12
g_disposicion_basural	0,597	0,585	0,610	0,576	0,965	0,04
g_lote_terreno_vivil	0,852	0,842	0,862	0,827	0,970	0,07
g_unipersonal1	0,124	0,118	0,131	0,154	1,242	0,08
j_femenino1	0,384	0,375	0,393	0,410	1,067	0,05
g_j_estado_civil1	0,061	0,056	0,066	0,054	0,882	0,03
j_sabe_leer1	0,935	0,930	0,939	0,959	1,026	0,12
g_j_grupo_empleo1	0,785	0,777	0,793	0,803	1,022	0,04
v_tiene_comp_notebook1	0,258	0,248	0,268	0,313	1,211	0,12
v_tiene_tableta1	0,076	0,071	0,083	0,125	1,631	0,15
v_tiene_aire1	0,590	0,578	0,602	0,578	0,979	0,02
v_tiene_microondas1	0,345	0,334	0,355	0,397	1,150	0,11
v_tiene_automovill	0,412	0,401	0,424	0,453	1,100	0,08
v_tiene_motocicleta1	0,541	0,530	0,552	0,533	0,985	0,02
zona_rural	0,384	0,374	0,394	0,308	0,802	0,16
v_dormitorios	2,264	2,242	2,286	2,217	0,980	0,04
d_personas	3,532	3,496	3,568	3,424	0,969	0,06
j_edad	49,534	49,136	49,933	48,350	0,976	0,07
j_anios_estudio	9,073	8,962	9,184	9,237	1,018	0,03
d_mujeres	1,797	1,775	1,819	1,718	0,956	0,07
d_menores10anios	0,608	0,590	0,626	0,559	0,919	0,06
d_edad10a25	0,919	0,898	0,941	0,863	0,938	0,06
d_edad25a65	1,721	1,702	1,740	1,730	1,005	0,01
d_mayores65anios	0,283	0,272	0,295	0,272	0,960	0,02
d_nro_remunerados	1,687	1,667	1,706	1,924	1,141	0,19

Notas: Para el cálculo de las medias de las variables del Censo, se recodificaron como valores perdidos (missing) aquellos datos no reportados en la base. Con el objetivo de asegurar la comparabilidad entre la base del CNPV 2022 y la EPHC 2023, se excluyeron de la base censal los registros correspondientes a población indígena, población en viviendas colectivas, personas sin vivienda, así como los departamentos de Alto Paraguay y Boquerón. Las diferencias de medias estandarizadas, es decir, la *d* de Cohen, miden la magnitud de la diferencia entre dos medias en términos de desviaciones estándar. Por ejemplo, una diferencia de medias estandarizada de 0,5 indica que la diferencia entre las medias es una desviación estándar de 0,5. A diferencia del valor *p*, que indica si las diferencias son estadísticamente significativas, la *d* de Cohen mide la magnitud de la diferencia. Como regla general, una *d* de Cohen inferior a 0,2 se considera una magnitud pequeña.

Al excluir estas tres variables, el universo de variables candidatas para el modelo se basó inicialmente en 22 variables (ver Anexo A) que cumplían con los tres criterios. Se construyeron 62 variables para capturar las características del hogar, características del jefe, propiedad de

activos, demografía del hogar y empleo del jefe de hogar, y que finalmente fueron parte del universo de variables candidatas para el modelo final.

3. Método de estimación

El método de estimación en áreas pequeñas utilizado en este estudio se basa en un principio fundamental: el bienestar de los hogares puede predecirse utilizando sus características observables, como educación del jefe de hogar, tipo de vivienda, acceso a servicios y composición demográfica.

En términos simples, el proceso funciona en dos etapas. Primero, se utiliza la EPHC 2023 para estimar un modelo estadístico que relaciona las características observables de los hogares con sus niveles de ingreso. Segundo, se aplican los parámetros estimados a todos los hogares del CNPV 2022 para predecir sus ingresos, para luego clasificarlos como pobres o no pobres.

El modelo reconoce que existen factores no observados que afectan el bienestar tanto a nivel de área geográfica como de hogar. Los efectos de área capturan factores locales como calidad de instituciones o acceso a mercados que afectan a todos los hogares de un territorio. Los efectos de hogar representan características específicas no observadas como redes sociales o salud familiar.

El mapa de pobreza distrital/barrial para Paraguay fue estimado a través del método de SAE, el cual está basado en el modelo originalmente propuesto por Battese et al. (1988), y extendido a modelo de error anidado por Elbers et al. (2003), y más recientemente Molina and Rao (2010).

El modelo se basa en el supuesto de que el bienestar de cada hogar (ya sea ingreso o consumo per cápita) en cada área está relacionado linealmente a un vector de características de ese hogar, de acuerdo con el siguiente modelo de error anidado:

$$y_{ch} = x_{ch}\beta + \eta_c + e_{ch}, \quad h = 1, \dots, N_c \quad c = 1, \dots, C,$$

en donde los términos de error η_c y e_{ch} representan, respectivamente, los errores a nivel de área y los errores específicos de cada hogar. Se asume que los mismos son independientes entre sí y siguen distribuciones normales, tal que:

$$\eta_c \text{ iid } \sim N(0, \sigma_\eta^2), e_{ch} \text{ iid } \sim N(0, \sigma_e^2),$$

donde las varianzas σ_η^2 y σ_e^2 son desconocidas. Aquí, C es el número de áreas administrativas en las que se divide la población y N_c es el número de hogares en el área C de la muestra de la encuesta.

Para la estimación del mapa de pobreza de Paraguay, se empleó la implementación de Molina y Rao (2010) propuesta por Corral et al. (2021). Se ajustó el modelo utilizando el método de Mínimos Cuadrados Generalizados (GLS) propuesto por Van der Weide (2014) obteniendo las estimaciones de los parámetros para la muestra observada:

$$\hat{\theta}_0 = (\hat{\beta}_0, \hat{\sigma}_\eta^2, \hat{\sigma}_e^2).$$

Estos parámetros luego se utilizaron en el censo para simular M vectores de bienestar, a través de simulaciones de Monte Carlo, procediendo de la siguiente manera:

$$y_{ch,r}^* = x_{ch,r} \hat{\beta}_0 + \eta_c^* + e_{ch,r}^*.$$

donde r representa el número de simulaciones, el término $\hat{\beta}_0$ se mantiene fijo para todos los M vectores siguiendo a Van der Weide (2014), y el término η_c^* se genera para las áreas que están en el censo y en la encuesta de la siguiente manera:

$$\begin{aligned} \eta_c^* &\sim N(\hat{\eta}_{c0}, \widehat{\text{var}}[\hat{\eta}_{c0}]), \\ \hat{\eta}_{c0} &= \hat{\gamma}_c \left(\sum_h \left(\frac{w_{ch}}{\hat{\sigma}_{e_0}^2} \right) \hat{e}_{ch} \right) \left(\sum_h \left(\frac{w_{ch}}{\hat{\sigma}_{e_0}^2} \right) \right)^{-1}, \\ \hat{\gamma}_c &= \frac{\hat{\sigma}_{\eta_0}^2}{\hat{\sigma}_{\eta_0}^2 + \sum_h w_{ch}^2 \left(\sum_h w_{ch} \sum_h \frac{w_{ch}}{\hat{\sigma}_{e_{cho}}^2} \right)^{-1}}, \\ \widehat{\text{var}}[\eta_c^*] &= \hat{\sigma}_{\eta_0}^2 - \hat{\gamma}_c^2 \left(\hat{\sigma}_{\eta_0}^2 + \sum_h \left(\frac{w_{ch}}{\hat{\sigma}_{e_{cho}}^2} \right)^2 \hat{\sigma}_{e_{cho}}^2 \right). \end{aligned}$$

Luego, los residuos específicos de cada hogar se obtienen a partir de:

$$e_{ch}^* = N(0, \hat{\sigma}_{e_{cho}}^2).$$

Así, el mejor estimador empírico del censo (Census EB) para cada área se obtiene al promediar las simulaciones de Monte Carlo de la siguiente forma:

$$\hat{\tau}_c^{EB} = \frac{1}{M} \sum_{m=1}^M \hat{\tau}_{cm}^*$$

donde τ es la tasa de incidencia de la pobreza (FGT0).

Para estimar la precisión de los estimadores, la metodología emplea un *bootstrap* paramétrico siguiendo a González-Manteiga et al. (2008) y a Corral et al. (2021). Todos los modelos y estimaciones se realizaron utilizando el paquete SAE más reciente para Stata desarrollado por Nguyen et al. (2018), empleando la última versión del software disponible al 22 de agosto de 2024.⁶

4. Desarrollo del modelo

La implementación del método SAE requiere múltiples decisiones metodológicas que afectan la calidad y precisión de las estimaciones finales. Esta sección documenta el proceso sistemático de selección del modelo empleado para el mapa de pobreza de Paraguay 2022, incluyendo la definición del nivel geográfico de los efectos de área, la transformación de la variable dependiente para cumplir los supuestos de normalidad, la elección entre modelos nacionales y subnacionales, los criterios de selección de variables explicativas, y la validación del cumplimiento de supuestos estadísticos. Cada decisión se fundamenta en criterios técnicos y se evalúa mediante pruebas estadísticas apropiadas para garantizar la robustez de los resultados.

4.1 Selección del nivel de los efectos por área

El Paraguay se divide en 17 departamentos y un distrito capital, la ciudad de Asunción. Los departamentos se dividen a su vez en distritos y luego en municipios, mientras que el distrito capital de Asunción se divide en barrios. La EPHC 2023 es representativa de la capital Asunción y de 15 departamentos, excluyendo Boquerón y Alto Paraguay en la zona norte.

El nivel geográfico más bajo en el que la EPHC 2023 pudo coincidir con el CNPV 2022 fue a nivel de distrito dentro de los departamentos y a nivel de barrio dentro de la capital. Por

⁶ El paquete de Stata más reciente se puede descargar de <https://github.com/pcorralrodas/SAE-Stata-Package>

esta razón, la estimación de tasas de pobreza monetaria en áreas pequeñas se realizó a nivel de distritos/barrios. No se dispuso de información estandarizada sobre municipios o unidades primaria de muestreo (UPM) a un nivel inferior al de los distritos comunes tanto en la EPHC 2023 como en el CNPV 2022 para explorar coincidencias por niveles administrativos más bajos (es decir, áreas debajo del nivel de distrito por departamentos y barrios representativos dentro del distrito capital de Asunción).

La Tabla 6 a continuación muestra el número de áreas (distritos/barrios) por departamento y el distrito capitalino de Asunción considerados en la construcción del mapa de pobreza. Como se observa, la menor cantidad de distritos se encuentra en los departamentos de Boquerón y Alto Paraguay, cada uno con 4 distritos. En tanto que el distrito de Asunción es el que presenta una mayor cantidad de áreas, con 67 barrios poblados, seguida por el departamento de Itapúa, con 30 distritos.

Tabla 6 - Número de Áreas (distritos/barrios) por departamento y el distrito capital Asunción

Código Departamento	Departamento	Nro. de Áreas (Distritos/Barrios)
0	Asunción (Barrios)	67
1	Concepción	14
2	San Pedro	22
3	Cordillera	20
4	Guairá	18
5	Caaguazú	22
6	Caazapá	11
7	Itapúa	30
8	Misiones	10
9	Paraguarí	18
10	Alto Paraná	22
11	Central	19
12	Ñeembucú	16
13	Amambay	6
14	Canindeyú	16
15	Presidente Hayes	10
16	Boquerón	4
17	Alto Paraguay	4
Total		329

Nota: Asunción cuenta con un total de 68 barrios, de los cuales, un barrio (Ñu Guasú) es una zona de recreación y no tiene hogares ni habitantes, por lo que no fue incluido en la construcción del mapa de pobreza.

La Tabla 7 a continuación presenta estadísticas sobre el número de hogares por área tanto en el censo como en la muestra. Por ejemplo, el distrito más pequeño (San Carlos del Apa, Concepción) alberga a 270 hogares, en tanto que los demás distritos albergan a más de 330 hogares en el CNPV 2022. El tamaño promedio de los distritos es de 5.269 hogares en el censo, en tanto que en la muestra es de 52 hogares. Así también, los resultados reflejan que el 25% de los distritos con menor número de hogares tiene menos de 1.416 hogares en el censo, en tanto que en la muestra tiene menos de 12 hogares. Por otra parte, el distrito mediano en el censo tiene 2.341 hogares y en la muestra 29 hogares. Además, el 25% de los distritos con mayor número de hogares tiene más de 4.722 hogares en el censo y 56 hogares en la muestra. El distrito más grande en el censo (Ciudad del Este, Alto Paraná) alberga a 92.313 hogares.

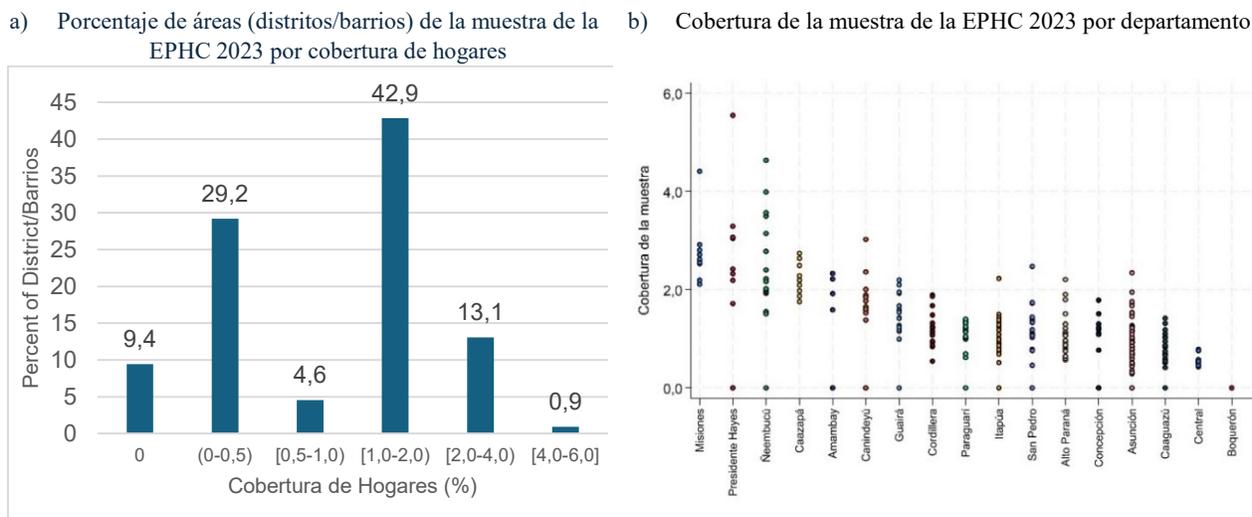
Tabla 7 - Resumen estadístico del número de hogares para 329 distritos y barrios del Paraguay

Estadístico	Número de hogares (Censo 2022)	Número de hogares (EPHC 2023)
Mínimo	270	0
Percentil 25	1.416	12
Mediana	2.341	29
Percentil 75	4.722	56
Media	5.269	52
Máximo	92.313	739

Figura 4-1 a continuación muestra el porcentaje de áreas (distritos/barrios) de la EPHC 2023 por cobertura de hogares. Los resultados de este análisis denotan que el 9,4% de las 329 áreas pobladas no estaban cubiertas por la EPHC 2023. El 29,2% de las áreas estaban cubiertas por menos del 0,5% de los hogares, en tanto que el 4,6% de las áreas cubrían entre el 0,5% y el 1% del total de hogares en las áreas. El 42,9% de las áreas estaban cubiertas por el 1-2% de la EPHC 2023. En tanto que el 13,1% de las áreas estaban cubiertas por el 2-4% y poco menos del 1% estaba cubierto por el 4 a 6% de los hogares. El panel b de la misma figura denota la cobertura por áreas, es decir, distritos por departamentos y barrios en Asunción, ordenado por cobertura. Así, la figura refleja el número de observaciones cubiertas por área, lo que puede ayudar a determinar la calidad del Censo EB derivado de la encuesta. Los resultados denotan que Misiones, Presidente Hayes y Ñeembucú se encuentran entre los departamentos con más áreas

que se beneficiarán con EB. En adición a estas figuras, la tabla en el Anexo B de este documento denota otras estadísticas descriptivas por departamento.

Figura 4-1 – Indicadores de cobertura de la EPHC 2023



Nota: Las estimaciones para los departamentos de Alto Paraguay y Boquerón son sintéticas, ya que estas áreas no fueron incluidas en la muestra de la EPHC 2023. Por tanto, sus valores se derivan exclusivamente del modelo aplicado sobre variables censales, sin información directa de encuesta. Estos resultados deben interpretarse con cautela.

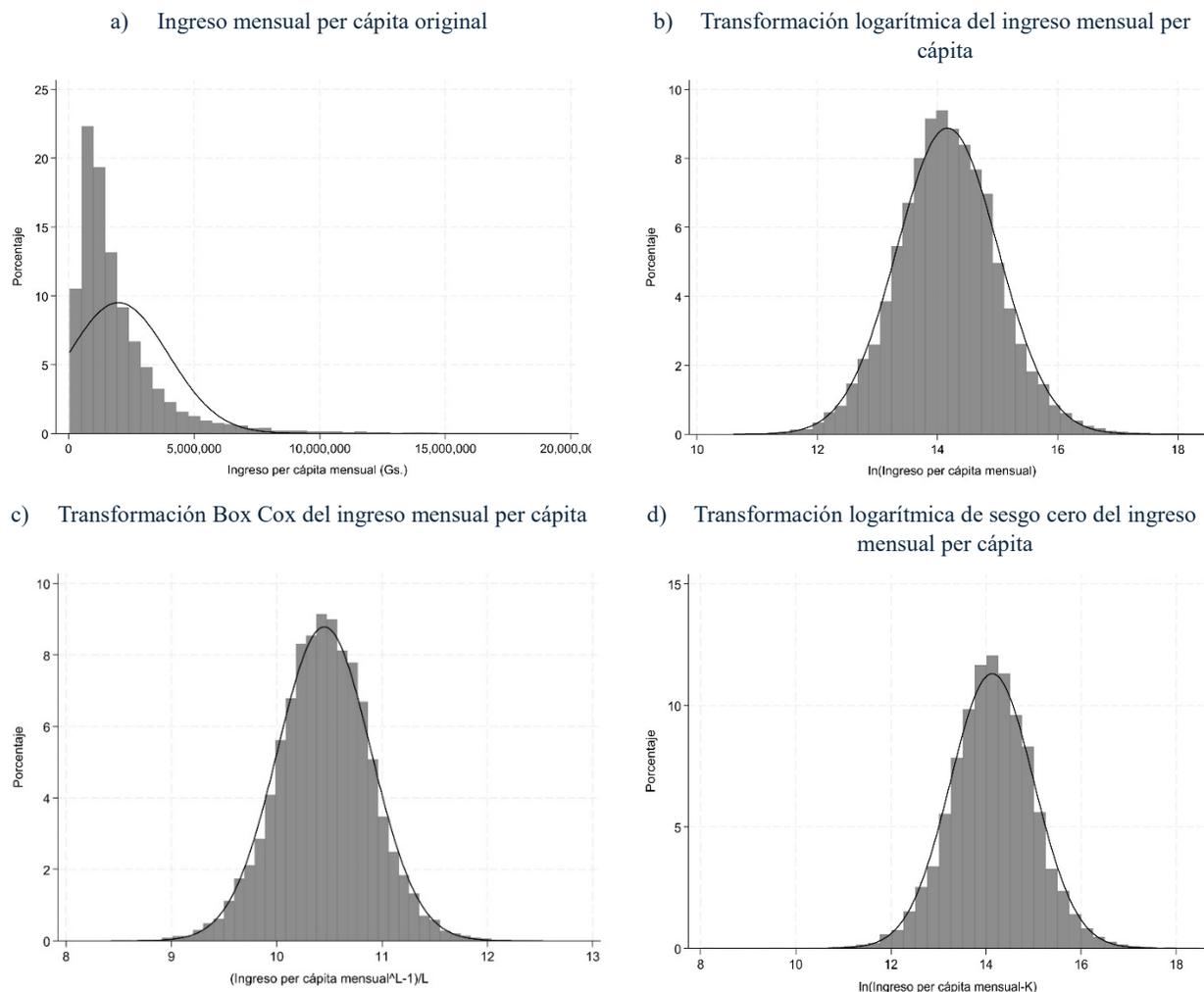
4.2 Transformación de la variable dependiente

Uno de los principales supuestos del modelo empleado en la construcción de mapas de pobreza es que los errores se distribuyen normalmente. Una variable dependiente distribuida normalmente a menudo hace que sea más probable que sus residuos también se distribuyan normalmente, por lo que es una condición deseable, aunque no estrictamente requerida, para modelos de regresión en general. Considerando que el modelo estimación EB toma a la variable de bienestar, en este caso el ingreso per cápita mensual del hogar, como variable dependiente, entonces es deseable que la misma presente una distribución normal. Sin embargo, el ingreso per cápita mensual en Paraguay no presenta una distribución normal, como se observa en el panel (a) de la Figura 4-2. La figura muestra que la distribución de los valores originales del ingreso per cápita mensual están más bien sesgados hacia la izquierda. Considerando que la transformación de variables para que se cumplan los supuestos subyacentes del modelo es uno de los aspectos principales de la estimación de áreas pequeñas basada en modelos, se consideraron diferentes transformaciones del agregado de bienestar (Figura 4-2, paneles b a d).

Para determinar qué transformación utilizar para que el agregado de bienestar pueda aproximarse a la normalidad, se evaluaron la asimetría y la curtosis de las distribuciones (típicamente una distribución normal se caracteriza por una asimetría de 0 y una curtosis de 3) y se realizaron pruebas estadísticas formales. Es importante destacar que las pruebas de normalidad formales son sensibles a tamaños de muestra grandes (generalmente de más de 2,000 observaciones) y en muestras más grandes, como en el conjunto de datos utilizados en este estudio, desviaciones menores pueden llevar a valores p significativos que conllevan al rechazo de la hipótesis nula de normalidad. Por ello, estas pruebas son sólo aplicables a tamaños de muestra más pequeños. No obstante, las mismas pueden ser informativas, y combinadas con otras métricas, como la asimetría y la curtosis, y el análisis gráfico pueden respaldar las decisiones sobre la determinación de las transformaciones que mejor imitan una distribución normal. En particular, se evaluaron tres alternativas de transformación: logarítmica, log de sesgo cero, y Box–Cox.

La figura muestra que la distribución de los valores originales del ingreso per cápita mensual están más bien sesgados hacia la izquierda. Por ello, en base a lo expuesto, y de manera a tener un mejor ajuste del modelo, se procedió a transformar los datos del ingreso a través de una transformación logarítmica (Panel (b), Figura 4-2), una transformación Box-Cox (Panel (c), Figura 4-2), y a través de una transformación logarítmica de sesgo cero (Panel (d), Figura 4-2), ya que todas contribuyen a normalizar la variable de bienestar, y, por tanto, a que los errores cumplan con los supuestos de normalidad del modelo.

Figura 4-2 - Comparación de distribuciones del ingreso per cápita mensual con diferentes transformaciones del ingreso per cápita



Las figuras denotan que la transformación Box-Cox (panel c) presenta un ajuste un poco mejor en el centro de la distribución relativo a la transformación logarítmica. Así también, las medidas de asimetría (0,00002) y curtosis (3,493) sugieren que esta transformación conlleva a que el ingreso per cápita presente una distribución más cercana a la normal. Por otra parte, los resultados de la prueba de Shapiro-Wilk para datos normales considerando la transformación de Box-Cox denota un valor p cercano a 0 por lo que no se rechaza la hipótesis nula de normalidad a niveles de significancia convencionales (ver Anexo D). Así también, el gráfico de cuantiles de la transformación de Box-Cox del ingreso per cápita frente a los cuantiles de la distribución normal, aunque no es perfecto, muestra una mayor concordancia con la normalidad que otras transformaciones consideradas (ver Anexos C a E). Las desviaciones en las colas

indican la presencia de curtosis, o más bien, la presencia de valores más extremos de lo esperado en una distribución normal. No obstante, la transformación de Box-Cox parece simular razonablemente una distribución normal, de modo a que se cumplan razonablemente los supuestos del modelo.

En resumen, la prueba con la transformación logarítmica tradicional mostró limitaciones, en particular por la presencia de valores muy bajos o cercanos a cero en el ingreso. El log de sesgo cero mejoró parcialmente este ajuste, pero no logró corregir de manera suficiente la asimetría. Finalmente, la transformación Box–Cox ofreció la mejor aproximación a la normalidad de la variable dependiente y permitió modelar de forma más adecuada la varianza de los residuos. En consecuencia, y considerando la premisa establecida por Corral et al. (2021), quienes sustentan que una transformación adecuada de los datos puede resultar en una reducción significativa de los errores cuadráticos medios, la transformación Box–Cox fue adoptada en la especificación final del modelo, al ofrecer un balance entre robustez estadística y facilidad de interpretación de los resultados.

4.3 Selección de un modelo nacional o varios modelos subnacionales

En esta sección se evaluó la conveniencia de estimar un único modelo nacional frente a la opción de construir modelos diferenciados para subpoblaciones específicas (por ejemplo, por región geográfica o por condición urbana/rural). La estimación de modelos separados puede, en principio, capturar mejor las peculiaridades locales, pero enfrenta limitaciones derivadas del tamaño muestral y la pérdida de eficiencia en la estimación.

En el caso paraguayo, la estimación de modelos subnacionales mostró problemas de estabilidad en los coeficientes y reducciones significativas en la precisión estadística. En contraste, el modelo nacional permitió aprovechar el mayor tamaño de muestra y conservar un conjunto consistente de variables explicativas. Además, el modelo nacional mostró un mejor desempeño en términos de ajuste global y capacidad predictiva.

Para explorar la selección del nivel del modelo, se realizaron varias simulaciones estadísticas en la muestra, la EPHC 2023. En particular, se muestreó aleatoriamente el 80% de la EPHC 2023 y se designó a este conjunto de datos como datos de entrenamiento para crear modelos estadísticos y obtener sus parámetros. El 20% de la muestra restante se utilizó como

datos de prueba para evaluar el rendimiento del modelo con datos no entrenados. Esto se realizó para explorar varias opciones de modelado, a saber:

- Un modelo nacional
- Un modelo rural y un modelo urbano
- Modelos para cada departamento
- Un modelo de la ciudad capital de Asunción, un modelo rural y un modelo urbano excluyendo Asunción.

La simulación se realizó 100 veces para cada uno de los modelos considerados y en cada ocasión se calculó el error cuadrático medio (ECM) para evaluar las predicciones del modelo en el conjunto de prueba. La Tabla 8 a continuación muestra el ECM para cada modelo simulado.

Tabla 8 – Comparativo de desempeño de modelos alternativos en datos de entrenamiento

Modelo	ECM [Intervalo de confianza del 95%]
Nacional	0,5510 [0,5488; 0,5532]
Rural/Urbano	0,5428 [0,5408; 0,5449]
Departamento	0,5482 [0,5462; 0,5504]
Rural/Urbano, Asunción	0,5434 [0,5412; 0,5455]

Nota: Basado en 100 simulaciones de selección aleatoria del 80% de la muestra para los datos de entrenamiento y el 20% de los datos de prueba.

Como se puede observar, basándose en el rendimiento del modelo, el modelo rural/urbano presenta una mayor probabilidad de predecir mejor los ingresos del hogar con un error cuadrático medio menor. Esto se observó tanto en la evaluación a nivel nacional como en la de Asunción (véase Anexo F de este documento). Además, es importante denotar que el mapa de distritos de Paraguay no delimita las zonas rurales y urbanas por dentro de los distritos. Por lo tanto, nuestro modelo incluyó un conjunto completo de parámetros que varían según la zona rural para permitir incorporar información sobre la relación entre la pobreza y las covariables propias de las zonas rurales. Sin embargo, las métricas de pobreza y las medidas de incertidumbre no se diferencian entre las zonas rurales y urbanas dentro de los distritos.

4.4 Selección de las variables a incluir en el modelo

La selección de variables para el modelo final se realizó mediante un proceso iterativo y sistemático. Inicialmente, se consideró el universo de variables de la EPHC 2023 que eran similares a las del censo, incluyendo promedios por departamento, distritos y barrios dentro de

Asunción. Se incluyó un indicador regional (área rural o urbana) y todas las variables a nivel de hogar se interactuaron con este indicador para así capturar las diferencias en las relaciones entre ingresos del hogar según área de residencia. El universo de variables incluyó un total de 87 variables candidatas, de las cuales, 62 variables fueron derivadas de las 22 inicialmente consideradas e incluyen 25 variables a nivel de hogar, 1 indicador rural, 9 variables a nivel de departamento y 27 variables a nivel de distrito/barrio. Las 25 variables a nivel de hogar varían además para el área rural, por lo que son adicionadas a las 62, totalizando así 87 variables (Ver Anexo G). Además, se descartaron 10 observaciones con un ingreso mensual per cápita del hogar de 0, y se eliminaron otras 8 observaciones con información incompleta sobre los niveles de educación y las características de la vivienda.

En una primera etapa, se aplicó el método *Lasso* con validación cruzada de 10 submuestras de las variables candidatas para identificar las variables con mayor poder explicativo. Este proceso resultó en la selección de 38 variables, con un R^2 ajustado de 0,63. Es importante mencionar que el método *Lasso* descarta las variables con alta probabilidad de tener un coeficiente igual a cero, lo que influye en todo el proceso posterior de selección de variables. En segundo lugar, se implementó un proceso recursivo de eliminación de variables, descartando aquellas con menor significancia estadística (es decir, con valores p superiores a 0,05). Este paso fue complementado con el análisis del Factor de Inflación de la Varianza (FIV) para mitigar la multicolinealidad.⁷ Al final de esta etapa, se obtuvo un modelo con 26 variables, con un promedio FIV < 3 y un R^2 ajustado de 0,62.

4.5 Cumplimiento de los supuestos

Como se mencionó anteriormente, un supuesto del modelo EB es que los residuos siguen una distribución normal. La violación de este supuesto, o bien, la desviación de la normalidad entre los residuos de los hogares, tanto en la parte inferior como en la superior de la distribución, podría causar sesgo en las tasas de pobreza estimadas (Newhouse et al., 2022). Por ello, es importante identificar el cumplimiento de este supuesto, ya que el no cumplimiento conlleva a implementar diferentes acciones de corrección

⁷ En este paso se ha optado por eliminar las covariables con un FIV superior a 3, como se muestra en la sección 6.3.1.1. Ejemplo 1 de las Directrices para la Estimación de Áreas Pequeñas para el Mapeo de la Pobreza.

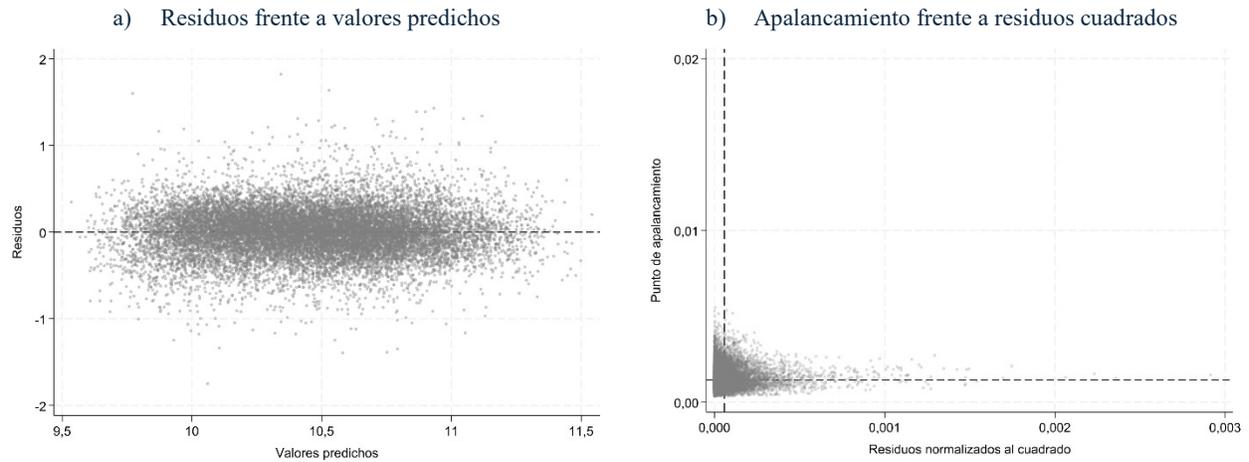
A lo largo del proceso de modelado, se tomaron varias medidas para garantizar que no se violaran los supuestos del modelo. Para asegurar el supuesto de linealidad (es decir, que la variable dependiente es una función lineal del vector de covariables seleccionado), se analizaron las relaciones entre la variable dependiente y las variables independientes mediante suavizadores polinómicos locales y análisis de residuos. En casos necesarios, las variables independientes se transformaron para aproximarse mejor a una relación lineal, por ejemplo, como en el caso del logaritmo natural del número de miembros del hogar. Además de la linealidad y la corrección de la multicolinealidad mediante la eliminación de covariables con altos factores de inflación de la varianza, como se mencionó anteriormente, también se realizó un análisis de influencia para detectar observaciones que podrían tener un impacto considerable en las estimaciones de los parámetros y, por lo tanto, en las predicciones del modelo. Esto incluye la identificación de valores atípicos (observaciones con residuos grandes) y observaciones influyentes, cuya omisión altera considerablemente las estimaciones de los parámetros. En esta etapa, se descartaron 62 observaciones clasificadas como valores extremos, quedando un total de 16.962 observaciones⁸.

El panel (a) de la Figura 4-3 muestra la gráfica de residuos frente a valores ajustados o predichos. Este gráfico sugiere que la nube de puntos es simétrica y no existen patrones que muestran evidencia de valores atípicos o incumplimiento de los supuestos de normalidad. A pesar de esto, se realizaron pruebas formales para comprobar la presencia de heterocedasticidad, específicamente, se aplicaron las pruebas de Breusch-Pagan y la de White. Los resultados de estas pruebas rechazan la hipótesis nula de homocedasticidad, es decir, la varianza constante de los residuos, lo que sugiere la posible presencia de heterocedasticidad (valor $p < 0,01$ en ambas pruebas), sugiriendo que el modelo alfa puede ayudar a reducir el sesgo y el ruido de las estimaciones⁹. Por otra parte, el panel (b) de la Figura 4-4 muestra la gráfica de apalancamiento frente a residuos al cuadrado tras descartar las observaciones influyentes (*outliers*), indicando que el cuadrante superior derecho de la gráfica, donde se ubicarían las observaciones influyentes, no presenta puntos que indiquen la persistencia de estas observaciones.

⁸ La identificación de observaciones problemáticas sigue varias reglas generales descritas en el apartado 6.3.2.1. Ejemplo 2, de las Directrices para la Estimación de Áreas Pequeñas para el Mapeo de la Pobreza, incluyendo alto apalancamiento, el estadístico de influencia D de Cook y los residuos estudentizados.

⁹ Véase apartado 4.2.3. El modelo Alfa de las Directrices para la estimación de áreas pequeñas para el mapeo de la pobreza.

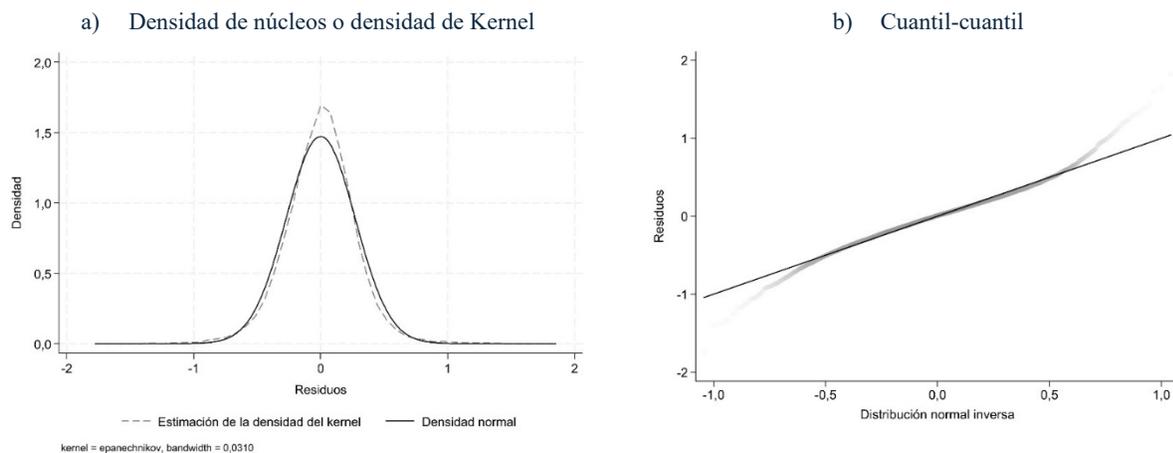
Figura 4-3 - Gráficos de residuos frente a valores predichos y apalancamiento frente a residuos cuadrados



Fuente: Elaboración propia con datos de la EPHC 2023 anual

Luego de remover los valores extremos, se analizaron dos gráficos más para evaluar el supuesto de normalidad. El panel (a) de la Figura 4-4 muestra el gráfico de densidad de Kernel, el cual indica que la distribución de residuos se aproxima a la normal, aunque no de forma perfecta. El panel (b), correspondiente al gráfico cuantil-cuantil, presenta en el eje horizontal los cuantiles de la distribución normal, y en el eje vertical los cuantiles de los residuos de la regresión. Este gráfico refleja una alineación de los residuos bastante cercana a la línea de 45 grados, con algunas desviaciones en los extremos, lo que se esperaría de una distribución normal.

Figura 4-4 - Gráfico de densidad de núcleos y gráfico cuantil-cuantil del ingreso per-cápita



Fuente: Elaboración propia con datos de la EPHC 2023 anual

Para investigar si estas ligeras desviaciones afectan los resultados del modelo, se procedió a comparar las predicciones del mismo con las estimaciones directas al nivel de representatividad de la encuesta, ya que estos resultados podrían resaltar cualquier sesgo vinculado a desviaciones de la normalidad. Para ello, se analizó el ajuste del modelo directamente a la encuesta de hogares y se compararon los niveles de ingresos de los hogares y las tasas de pobreza predichos con los observados, según el nivel de representatividad de la encuesta (véase Tabla 9). Estos resultados muestran una alta concordancia con las estimaciones directas, lo que garantiza que las desviaciones de la normalidad no sesgan los resultados de la predicción de la pobreza.

4.6 Evaluación de las estimaciones

El último paso en el proceso de modelado consiste en la validación del modelo, donde las estimaciones puntuales se comparan con parámetros de referencia conocidos. En este paso, el modelo se ajusta a los datos de la encuesta EPHC 2023, utilizados para desarrollarlo, y se compara el bienestar previsto con el observado según el nivel de representatividad de la encuesta para explorar cualquier posible sesgo relacionado con desviaciones de la normalidad. Los resultados de este análisis se presentan en la Tabla 9 a continuación, la cual compara las predicciones del modelo SAE aplicadas a la encuesta EPHC 2023 con los estimadores directos de pobreza de la EPHC 2023, revelando que todas las estimaciones de pobreza SAE caen dentro de los intervalos de confianza de los estimadores directos que utilizan la muestra de la encuesta.

Tabla 9 – Comparación de estimadores del modelo SAE vs. estimadores directos de la pobreza en la EPCH 2023

Departamento	Número de Hogares EPHC 2023	Ingreso Mensual per Cápita		Pobreza Total			
		Predicción SAE en la EPCH 2023	Est. Directa EPCH 2023	Predicción SAE en el EPCH 2023	Est. Directa EPCH 2023	Lim. Inf. 95% Est. Dir. EPCH 2023	Lim. Sup. 95% Est. Dir. EPCH 2023
Asunción	124.227	3.005.979	3.430.591	10,4%	9,7%	7,1%	13,1%
Concepción	57.299	1.233.469	1.356.244	37,0%	35,9%	30,0%	42,2%
San Pedro	105.601	1.282.110	1.429.946	34,2%	32,4%	28,4%	36,8%
Cordillera	75.168	1.536.014	1.533.825	25,2%	23,9%	19,2%	29,4%
Guairá	59.518	1.309.350	1.261.862	33,3%	34,0%	27,9%	40,7%
Caaguazú	128.908	1.308.388	1.611.279	34,6%	31,9%	26,8%	37,4%
Caazapá	46.641	1.220.607	1.310.478	36,8%	36,8%	28,8%	45,6%
Itapúa	133.837	1.427.008	1.558.141	30,4%	29,8%	25,9%	34,0%
Misiones	33.045	1.558.093	1.741.892	26,6%	26,4%	22,2%	31,2%
Paraguarí	56.616	1.293.091	1.278.262	34,4%	30,5%	25,4%	36,0%
Alto Paraná	206.962	1.737.436	1.930.270	23,3%	23,3%	20,3%	26,6%
Central	495.23	2.064.740	2.038.905	14,9%	13,5%	11,7%	15,5%
Ñeembucú	24.884	1.614.201	1.693.032	2,3,3%	26,0%	20,3%	32,7%
Amambay	43.91	1.750.353	1.847.233	20,5%	22,4%	18,7%	26,5%
Canindeyú	50.201	1.487.217	1.689.922	31,1%	35,1%	28,8%	42,1%
Presidente Hayes	24.186	1.629.583	1.832.445	21,6%	18,8%	15,6%	22,5%
Boquerón							
Alto Paraguay				no muestreado			

Nota: Las estimaciones para los departamentos de Alto Paraguay y Boquerón son sintéticas, ya que estas áreas no fueron incluidas en la muestra de la EPCH 2023. Por tanto, sus valores se derivan exclusivamente del modelo aplicado sobre variables censales, sin información directa de encuesta. Estos resultados deben interpretarse con cautela.

5. Resultados del modelo

5.1 Selección del modelo final

El modelo final incorpora 21 variables explicativas y una constante, con un ajuste notable. Las variables seleccionadas representan diversos aspectos socioeconómicos de los hogares paraguayos, como las características de la vivienda, la posesión de bienes duraderos, los indicadores del hogar, la composición demográfica de los hogares, la ubicación geográfica (diferencias rurales) y los promedios por distritos y barrios¹⁰.

La Tabla 10 a continuación presenta los coeficientes estimados y las estadísticas asociadas del modelo econométrico final (Modelo Beta). En las zonas rurales, las viviendas más grandes (con más habitaciones por persona) y la posesión de motocicletas y automóviles mostraron una fuerte relación con los ingresos del hogar. A nivel de distrito/barrio, las zonas con un mayor porcentaje de hogares encabezados por mujeres tendían a ser más pobres, mientras que las zonas que reportaron mayores niveles de tenencia de tabletas o artículos similares, o que se negaban a responder preguntas sobre la propiedad de activos, se asocian con mayores ingresos del hogar.

En la potencial presencia de heterocedasticidad discutida en la Sección 4.5, se procedió a implementar el modelo alfa, ya que la aplicación de este modelo para obtener estimaciones de Census EB puede arrojar mejores resultados en términos de sesgo y error cuadrático medio (véase la sección 4.2.3 de las Directrices para la Estimación de Áreas Pequeñas para el Mapeo de la Pobreza). El modelo alfa consiste en la estimación de un modelo en donde la variable dependiente son los residuos del modelo base (Beta), y las variables independientes son aquellas que pasaron la selección inicial de *Lasso*, pero que no se incluyeron en el modelo de ingresos transformado. Se aplicó nuevamente la técnica de *Lasso*, siguiendo el mismo proceso de selección de variables para el modelo principal discutido en la sección 4, con eliminación recursiva y análisis FIV, obteniendo ocho variables significativas superiores al nivel de 0,01 y con FIV < 5. Los resultados del modelo alfa presentados en la Tabla 10 incluyó ocho variables significativas superiores al valor de significancia de 0,01 y un R² de 0,01

¹⁰ Siguiendo las mejores prácticas internacionales para estadísticas experimentales, los resultados se presentan por rangos para facilitar su correcta interpretación estadística, considerando los intervalos de confianza asociados a cada estimación.

Tabla 10 - Parámetros estimados del modelo final

Modelo Beta	Coefficiente		Error std.
g_agua_utilizada1	0,0349	***	(0,0069)
g_disposicion_basural	0,0661	***	(0,0076)
g_tipo_combustible1	0,1093	***	(0,0076)
g_tipo_pared1	0,0548	***	(0,0084)
g_tipo_techo1	0,0373	***	(0,0063)
v_tiene_aire1	0,1186	***	(0,0068)
v_tiene_automovill	0,1350	***	(0,0074)
v_tiene_comp_notebook1	0,1753	***	(0,0071)
v_tiene_microondas1	0,0983	***	(0,0065)
ln_d_personas	-0,1838	***	(0,0064)
d_porc_jovenes	-0,0724	***	(0,0128)
d_porc_mayores	0,1033	***	(0,0118)
d_porc_remunerados	0,5144	***	(0,0104)
j_feminino1	-0,0221	***	(0,0055)
j_sabe_leer1	0,0557	**	(0,0101)
rur_ln_v_dormitorios_pc	0,0276	**	(0,0101)
rur_v_tiene_automovill	0,0242	***	(0,0113)
rur_v_tiene_motocicleta1	0,0189	***	(0,0080)
distbar_j_feminino1	-0,3677	***	(0,0952)
distbar_perd_activo	0,6822	***	(0,1284)
distbar_v_tiene_tableta1	0,4794	***	(0,0451)
Constante	10,4033	***	(0,0396)
Número de observaciones	16.962		
R cuadrado ajustado	0.6175		
R cuadrado	0,6180		
Raíz cuadrada del error cuadrado medio	0,2927		
F estadístico	1.305,1611		
Modelo Alfa	Coefficiente		Error std.
ln_v_dormitorios_pc	0,2175	***	(0,0466)
v_tiene_motocicleta1	-0,1699	***	(0,0463)
rur_d_porc_jovenes	0,3677	***	(0,1249)
rur_g_disposicion_basural	-0,3338	***	(0,1209)
rur_g_lote_terreno_vivi1	0,2870	***	(0,0635)
rur_g_tipo_combustible1	-0,3355	***	(0,0681)
rur_j_anios_estudio	0,0245	***	(0,0084)
distbar_g_tipo_pared1	-0,6747	***	(0,1256)
Constante	-4,9270	***	(0,1164)
Número de observaciones	16.962		
R cuadrado ajustado	0,0138		

R cuadrado	0,0143
Raíz cuadrada del error cuadrado medio	2,3325
F estadístico	30,7276
Sigma eta cuadrado	0,0025
Sigma Eta cuadrado/MSE	0,0298
Varianza de épsilon	0,0833

Nota: *p<0.1; **p<0.05, *** p<0.01

5.2 Métricas de ajuste y precisión

El modelo final, cuenta con un R^2 ajustado de 0,618 después de una ronda final de eliminación recursiva, y comprende 21 variables explicativas, todas las cuales son significativas a un nivel de significancia del 1% (0,01), y ocho variables para explicar los residuos, también todas significativas más allá del nivel de significancia del 1% (véase Tabla 10 y Anexo G).

Adicionalmente, se reportaron métricas de precisión, incluyendo el error cuadrático medio (ECM) y los errores estándar por área, presentados en la Tabla 8 y en el Anexo F. Estas métricas muestran niveles adecuados de confiabilidad para los propósitos del mapeo de pobreza.

5.3 Comparación de los resultados con los estimadores directos

Con el objetivo de validar el modelo, se realizó un análisis comparativo de las estimaciones de tasas de pobreza en el censo con los estimadores directos de tasas de pobreza en la encuesta. Los resultados se presentan en la Tabla 11 a continuación. En general se observa que las tasas de pobreza estimadas por SAE tienden a ser más altas que las tasas de pobreza por estimadores directos, con algunas excepciones. Este resultado es el reflejo de estimaciones promedios de ingresos mensuales per cápita más bajos que lo observado en la encuesta de hogares. En cuatro departamentos (Itapúa, Alto Paraná, Canindeyú y Presidente Hayes), la estimación de la tasa de pobreza total del modelo SAE en el censo se encuentra fuera del intervalo de confianza de las estimaciones directas de la encuesta, aunque generalmente por menos de 2,5 puntos porcentuales del intervalo de confianza inferior. En el Anexo H se presentan los rangos de pobreza a nivel departamental a partir de estimaciones directas con rangos obtenidos mediante la agregación de los resultados de la SAE que muestran un nivel muy alto de correlación entre las estimaciones de la SAE y las estimaciones directas del EPHC y un rango consistente en los departamentos por niveles de ingreso y pobreza.

Tabla 11 - Predicción SAE en el Censo 2022 y estimadores directos en la EPHC 2023

Departamento	Número de Hogares Censo	Ingreso Per Cápita		Pobreza Total			
		Predicción SAE en el Censo	Est. Directa EPHC 2023	Predicción SAE en el Censo	Est. Directa EPHC 2023	Lim. Inf. 95% Est. Dir. EPHC 2023	Lim. Sup. 95% Est. Dir. EPHC 2023
Asunción	133.456	3.470.313	3.430.591	7,3%	9,7%	7,10%	13,10%
Concepción	55.895	1.431.967	1.356.244	31,1%	35,9%	30,00%	42,20%
San Pedro	104.147	1.260.534	1.429.946	34,2%	32,4%	28,40%	36,80%
Cordillera	79.107	1.567.196	1.533.825	23,7%	23,9%	19,20%	29,40%
Guairá	55.873	1.428.658	1.261.862	29,4%	34,0%	27,90%	40,70%
Caaguazú	123.318	1.489.092	1.611.279	29,9%	31,9%	26,80%	37,40%
Caazapá	42.056	1.285.180	1.310.478	34,7%	36,8%	28,80%	45,60%
Itapúa	140.21	1.648.015	1.558.141	24,5%	29,8%	25,90%	34,00%
Misiones	34.516	1.737.882	1.741.892	22,5%	26,4%	22,20%	31,20%
Paraguari	62.3	1.442.314	1.278.262	27,3%	30,5%	25,40%	36,00%
Alto Paraná	216.913	2.020.847	1.930.270	17,9%	23,3%	20,30%	26,60%
Central	514.841	2.225.579	2.038.905	14,7%	13,5%	11,70%	15,50%
Ñeembucú	26.667	1.710.688	1.693.032	21,2%	26,0%	20,30%	32,70%
Amambay	47.826	1.859.657	1.847.233	21,2%	22,4%	18,70%	26,50%
Canindeyú	52.578	1.643.925	1.689.922	26,5%	35,1%	28,80%	42,10%
Presidente Hayes	26.723	1.746.041	1.832.445	23,0%	18,8%	15,60%	22,50%
Boquerón*	13.344	2.521.658		8,3%			
Alto Paraguay*	3.889	1.315.524		38,7%		<i>no muestreado</i>	

Fuente: Elaboración propia con datos de la EPHC 2023 anual y el CNVP 2022.

Nota: *Las estimaciones para los departamentos de Alto Paraguay y Boquerón son sintéticas, ya que estas áreas no fueron incluidas en la muestra de la EPHC 2023. Por tanto, sus valores se derivan exclusivamente del modelo aplicado sobre variables censales, sin información directa de encuesta. Estos resultados deben interpretarse con cautela.

Para comprender las diferencias a nivel departamental entre las predicciones del censo y de la encuesta, se analizaron los promedios de las variables del modelo en ambos conjuntos de datos para cada departamento. La Tabla del Anexo I presenta, para cada departamento, la razón entre la media de la EPHC 2023 y la media del Censo 2022 de todas las variables incluidas en el modelo SAE, excepto las variables a nivel de área derivadas del Censo 2022. Con base a estas ratios, las diferencias en tasas de pobreza entre la predicción del SAE en el Censo 2022 en Itapúa parecen ser explicadas por las diferencias en la propiedad de activos declarada, con una mayor tenencia de computadoras o laptops y microondas en Itapúa en el censo. En Alto Paraguay, la propiedad de activos declarada en términos de computadoras o laptops y un mayor acceso declarado a fuentes de agua de mayor calidad utilizadas por el hogar, como (1) pozo artesiano, (2) ESSAP, (3) red o prestador privado, (4) pozo con bomba, o (5) SENASA o junta de

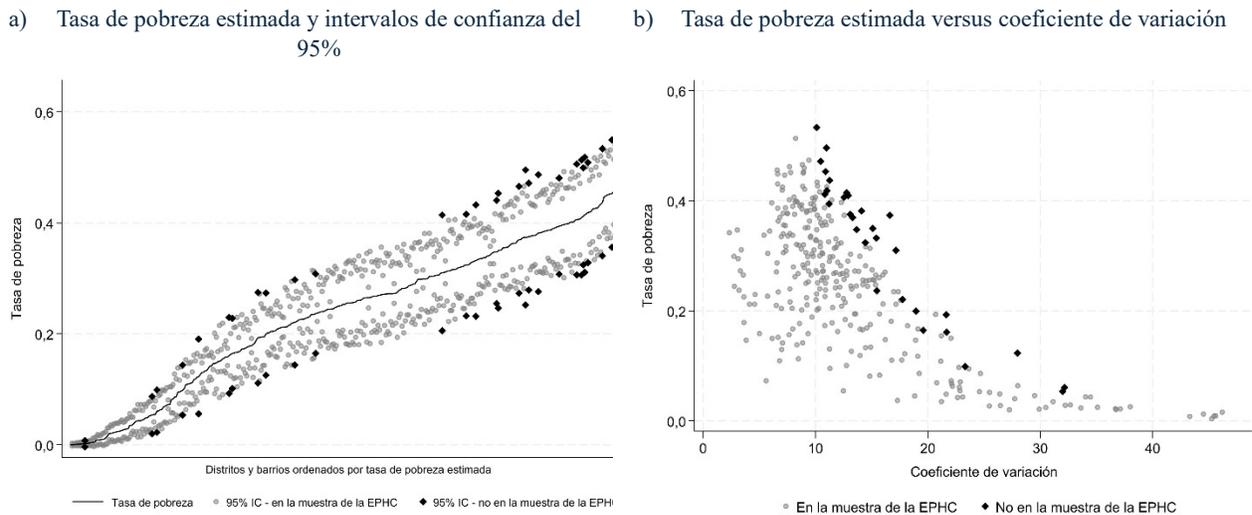
saneamiento, están conllevando a tasas de pobreza más bajas. Asimismo, un mayor acceso declarado a fuentes de agua de mayor calidad utilizadas por el hogar parece estar impulsando tasas de pobreza más bajas en Canindeyú. A pesar de estas diferencias, las tasas de pobreza estimadas están cerca de los intervalos de confianza y no parecen ser preocupaciones importantes.

Es importante señalar que la precisión de las estimaciones varía según la cobertura de la encuesta en cada área. Los departamentos incluidos en la muestra EPHC 2023 presentan intervalos de confianza menores, mientras que las estimaciones sintéticas para Alto Paraguay y Boquerón tienen mayor incertidumbre estadística, como es esperado en este tipo de metodologías.

5.4 Comparación entre las áreas muestreadas en la encuesta y las no muestreadas

De manera a comprender mejor la precisión de los resultados obtenidos por SAE, se confeccionaron dos gráficos. El primero, denotado en el panel a de la Figura 5-1, muestra las tasas de pobreza estimadas para las áreas muestreadas y no muestreadas en la EPHC 2023, junto con los intervalos de confianza del 95%. Esta figura denota que las áreas no incluidas en la EPHC 2023 tienden a mostrar intervalos de confianza más amplios, indicando estimaciones menos precisas que las áreas muestreadas. En tanto el panel b de la misma figura muestra el coeficiente de variación contra las tasas de pobreza para las áreas muestreadas y no muestreadas en la EPHC 2023, denotando que, en general, las áreas que no conformaron parte de la muestra, tienden a presentar un mayor coeficiente de variación que las áreas incluidas en la muestra.

Figura 5-1 – Comparativo entre áreas muestreadas en la encuesta y las no muestreadas



6. Mapa de pobreza

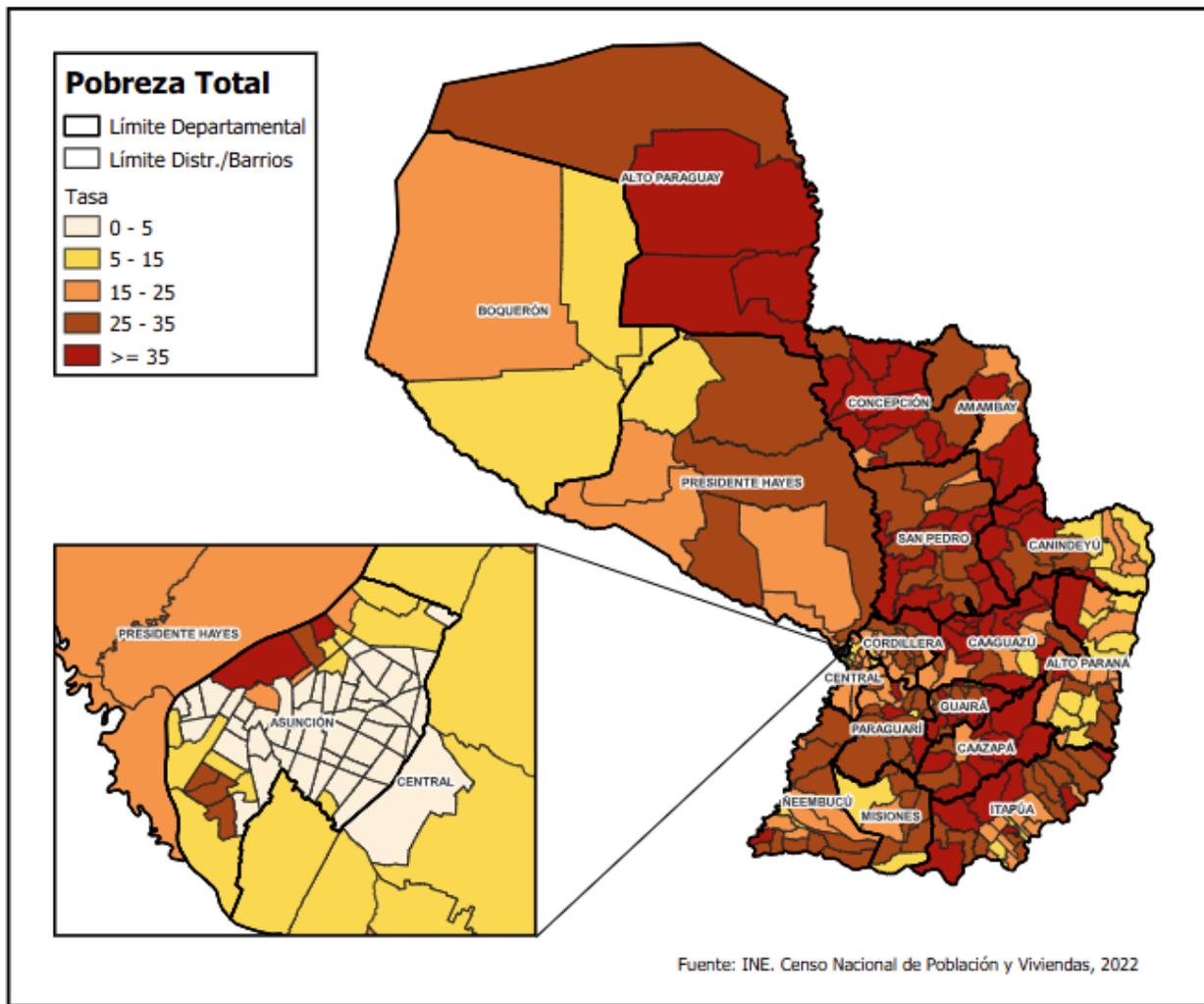
En esta sección se presenta el mapa de pobreza de Paraguay como también de la medida de bienestar en Paraguay, el ingreso per cápita del hogar. La examinación de estos mapas permite identificar las disparidades geográficas y los principales focos de concentración de la población pobre.

6.1 Mapa de pobreza utilizando el ingreso per cápita como medida de bienestar

La Figura 6-1 presenta el mapa de pobreza total para Paraguay, basado en la estimación de ingreso per cápita mensual obtenida con el método SAE y aplicando las líneas de pobreza oficiales. La distribución espacial de las tasas de pobreza denota una alta dispersión geográfica tanto a lo largo del país, como al interior de cada departamento y de la capital del país, Asunción. Notablemente, el departamento de mayor tasa de pobreza es Alto Paraguay (38,7%), en la región del Chaco, seguido por Caazapá (34,7%), San Pedro (34,2%), y Concepción (31,1%). En cuanto a distritos, el distrito con la tasa de pobreza más alta a nivel nacional es el distrito de Paso Barreto, situado en el departamento de Concepción, con una tasa de pobreza del 53,3%, seguido por Sargento José Félix López, con una tasa del 51,3%, en el mismo departamento. En contraste, el distrito con menor tasa de pobreza es Asunción, con una tasa de pobreza del 7,3%. Sin embargo, incluso dentro de Asunción existe una amplia heterogeneidad en las tasas de pobreza a

nivel de barrios. Por ejemplo, el barrio con la tasa más alta de pobreza es Banco San Miguel, con una tasa del 42,5%, seguido por el barrio Bañado Cara Cara, con una tasa del 39,7%. En tanto que el barrio con la tasa de menor pobreza es el barrio Mburicao, con 0,03%.

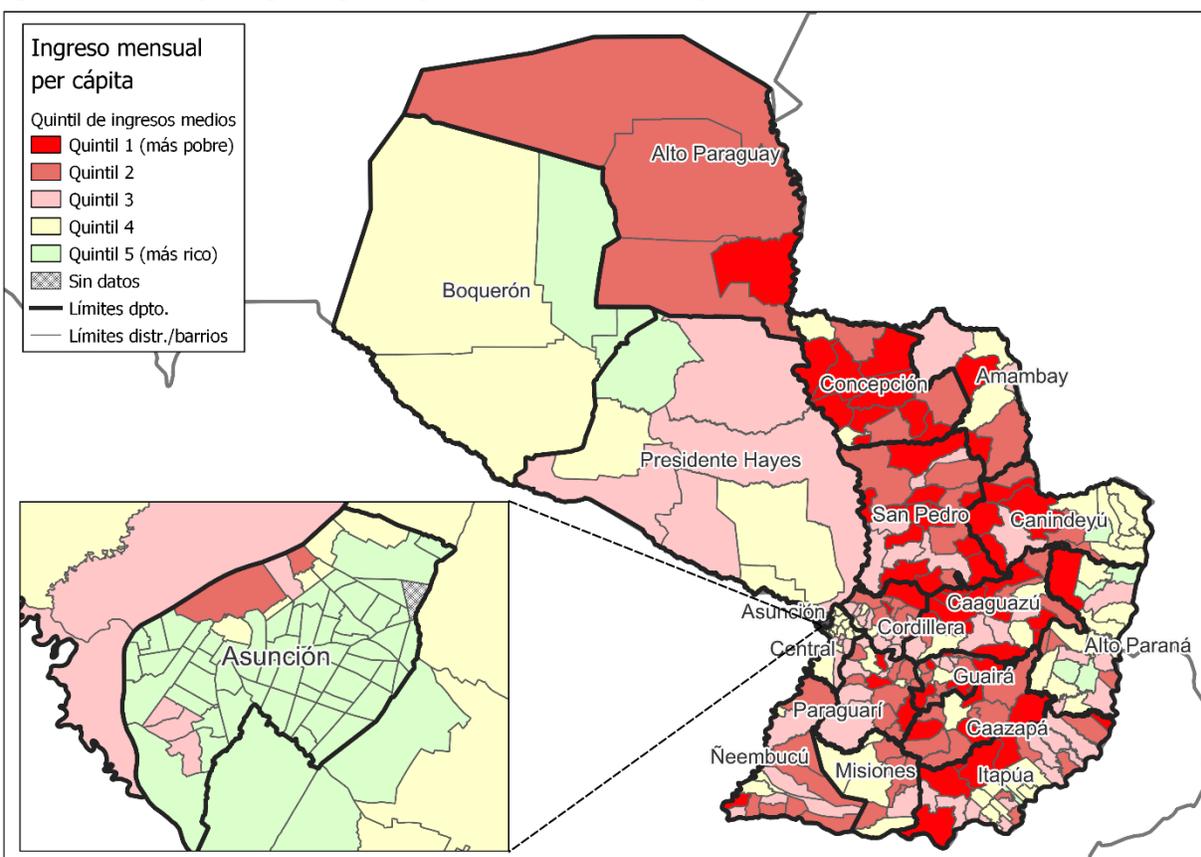
Figura 6-1 – Mapa de pobreza total distrital de Paraguay



Nota: Las estimaciones para Alto Paraguay y Boquerón son sintéticas (departamentos no incluidos en muestra EPHC 2023).

La Figura 6-2 refleja el ingreso per cápita mensual promedio por distrito, evidenciando una alta heterogeneidad en la distribución del ingreso. Con respecto a otros distritos del país, el ingreso per cápita mensual es más elevado en Asunción que en otros distritos del país.

Figura 6-2 – Mapa ingreso per cápita mensual de Paraguay



Nota: Las estimaciones para Alto Paraguay y Boquerón son sintéticas (departamentos no incluidos en muestra EPHC 2023).

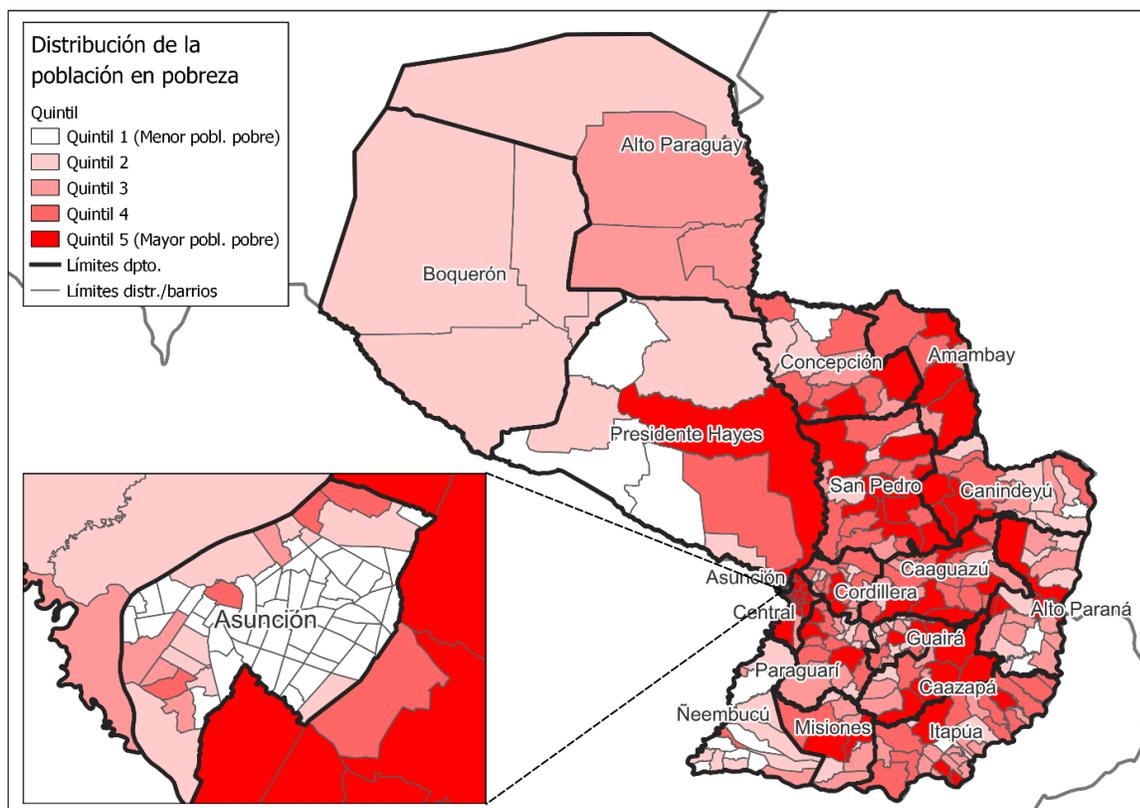
6.2 Distribución territorial de la población en pobreza

La Figura 6-3 presenta la distribución espacial de la población en situación de pobreza aplicando la metodología SAE a lo largo del territorio paraguayo. De los 5.928.804 habitantes no indígenas de Paraguay, se estima que 1.222.786 viven en condiciones de pobreza monetaria, representando una tasa de nacional del 20,6%.

La concentración geográfica de la pobreza revela patrones territoriales específicos que trascienden las divisiones administrativas tradicionales. El distrito de Ciudad del Este, en el departamento de Alto Paraná, concentra la mayor proporción de personas pobres del país (4,2%), seguido por Capiatá, en Central (2,9%), y en el mismo departamento por Luque (2,7%) y Caaguazú, en el departamento de Caaguazú (2,7%). Completan los principales focos de concentración absoluta de pobreza los distritos de Limpio (2,3%) y San Lorenzo (2,0%), todos ubicados en el departamento Central, y el distrito de Coronel Oviedo en Caaguazú (1,6%).

Esta distribución evidencia que, si bien las tasas de pobreza pueden ser menores en áreas urbanas densamente pobladas, el volumen absoluto de personas en situación de vulnerabilidad se concentra en centros urbanos y periurbanos de rápido crecimiento, planteando desafíos específicos para la política social.

Figura 6-3 – Mapa de la distribución de la población en pobreza



Nota: Las estimaciones para Alto Paraguay y Boquerón son sintéticas (departamentos no incluidos en muestra EPHC 2023).

7. Posibles extensiones

En esta sección se presentan posibles extensiones del mapa de pobreza para análisis territorial y focalización de políticas públicas. Los ejercicios aquí mostrados constituyen demostraciones metodológicas. Análisis más detallados serán desarrollados por el INE una vez se publiquen el Índice de Pobreza Multidimensional y otros indicadores complementarios a nivel distrital actualmente en preparación.

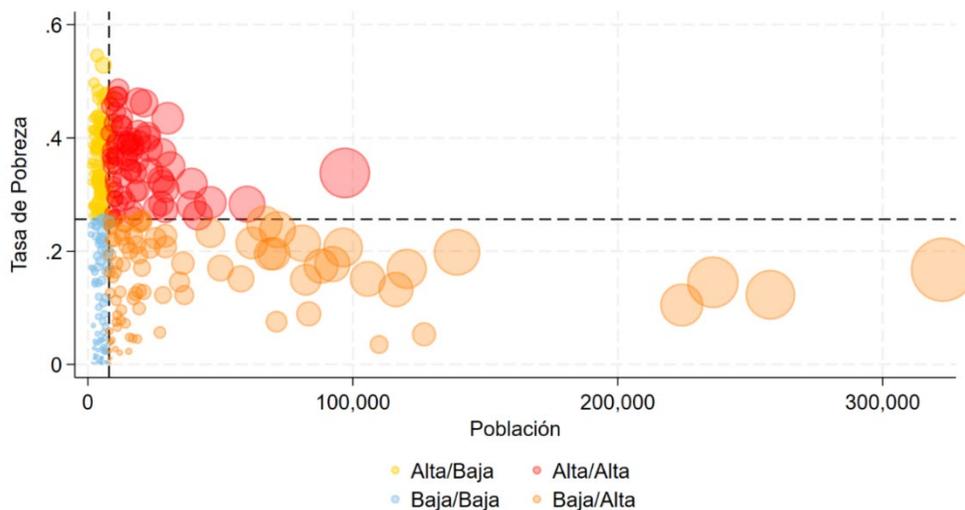
El análisis se estructura en tres partes. Primero, se analiza la distribución territorial de la pobreza desde una perspectiva de focalización de políticas públicas, desarrollando tipologías que

faciliten la priorización de intervenciones según combinaciones de incidencia de pobreza y tamaño poblacional. Segundo, se explora la relación entre pobreza monetaria y factores estructurales mediante el análisis comparativo con indicadores de necesidades básicas insatisfechas, proporcionando evidencia empírica sobre los determinantes territoriales de la pobreza. Finalmente, se discuten las extensiones metodológicas y temáticas que podrían enriquecer futuras actualizaciones del mapa de pobreza. Los resultados se interpretan considerando tanto su significancia estadística como su relevancia para el diseño de políticas públicas territorialmente diferenciadas.

7.1 Implicaciones para la focalización de políticas públicas

Los resultados de los mapas de pobreza proporcionan información crucial para el diseño de políticas públicas, y en particular, para los programas de focalización. Por ejemplo, la combinación de la incidencia de pobreza con el tamaño poblacional (a nivel distrital o de barrio) permite identificar tipologías territoriales que requieren enfoques de política distintos. La Figura 7-1 presenta un diagrama de dispersión que facilita la priorización y focalización de programas sociales, con cuatro cuadrantes divididos entre la mediana de la tasa de pobreza (25,6%) y la mediana del tamaño poblacional (7.533 habitantes):

Figura 7-1 - Diagrama de dispersión de cuatro cuadrantes de la tasa de pobreza y población



* El tamaño del círculo indica el tamaño de la población en pobreza.
Las líneas de referencia indican el valor mediano en el eje correspondiente.

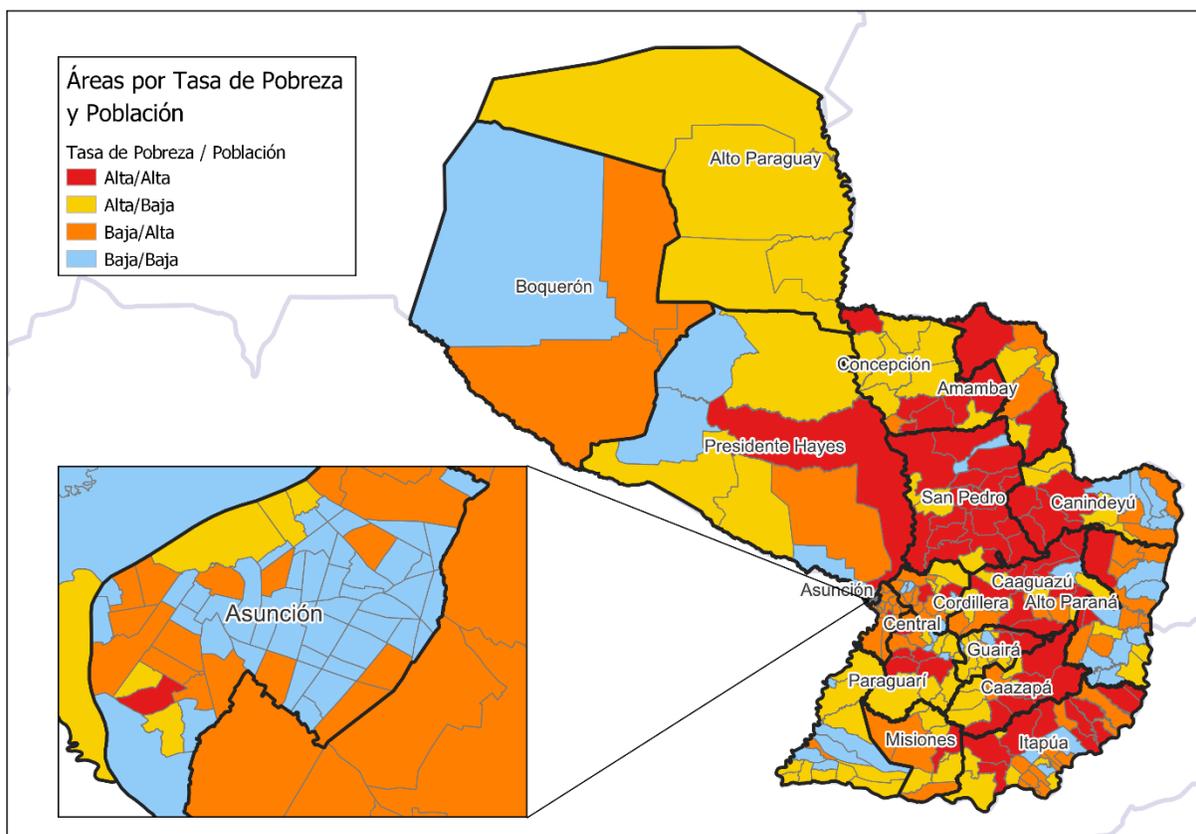
A partir de esta figura, se identifican cuatro tipologías que pueden guiar la priorización de intervenciones, como se denota en la tabla a continuación:

Tabla 12 – Tipologías de niveles de pobreza y densidad poblacional

Cuadrante superior izquierdo (Alta pobreza + Baja población)	Cuadrante superior derecho (Alta pobreza + Alta población)
<u>Enfoque en la equidad:</u> Programas comunitarios amplios adaptados a las características de baja densidad.	<u>Máxima prioridad:</u> Programas integrales y comprehensivos que aborden la intensidad y magnitud del problema.
Cuadrante inferior izquierdo (Baja pobreza + Baja población)	Cuadrante inferior derecho (Baja pobreza + Alta población)
<u>Prioridad relativa menor:</u> Monitoreo y programas específicos según el contexto.	<u>Enfoque en la eficiencia:</u> Programas específicos y bien focalizados hacia poblaciones vulnerables.

El análisis de focalización puede enriquecerse mediante la incorporación de la densidad poblacional como variable adicional, ya que esta medida proporciona información complementaria sobre la concentración espacial de la pobreza y las economías de escala potenciales para las intervenciones. La Figura 7-2 complementa este análisis al mostrar la distribución geográfica combinada de incidencia de pobreza y densidad poblacional. Los resultados sugieren que las áreas dentro y alrededor del departamento de San Pedro, junto con sectores del Chaco, presentan la combinación más crítica de altas tasas de pobreza y una alta densidad poblacional, constituyendo las áreas de máxima prioridad para intervenciones integrales. En contraste los centros urbanos del departamento Central, que muestran densidades poblacionales elevadas, pero tasas de pobreza relativamente menores, requieren programas específicos de reducción de pobreza dirigidos hacia bolsones de vulnerabilidad dentro de contextos urbanos más prósperos.

Figura 7-2– Mapa de distribución de incidencia de pobreza y densidad poblacional



Nota: Las estimaciones para Alto Paraguay y Boquerón son sintéticas (departamentos no incluidos en muestra EPHC 2023).

7.2 Análisis de factores estructurales: pobreza monetaria y necesidades básicas insatisfechas

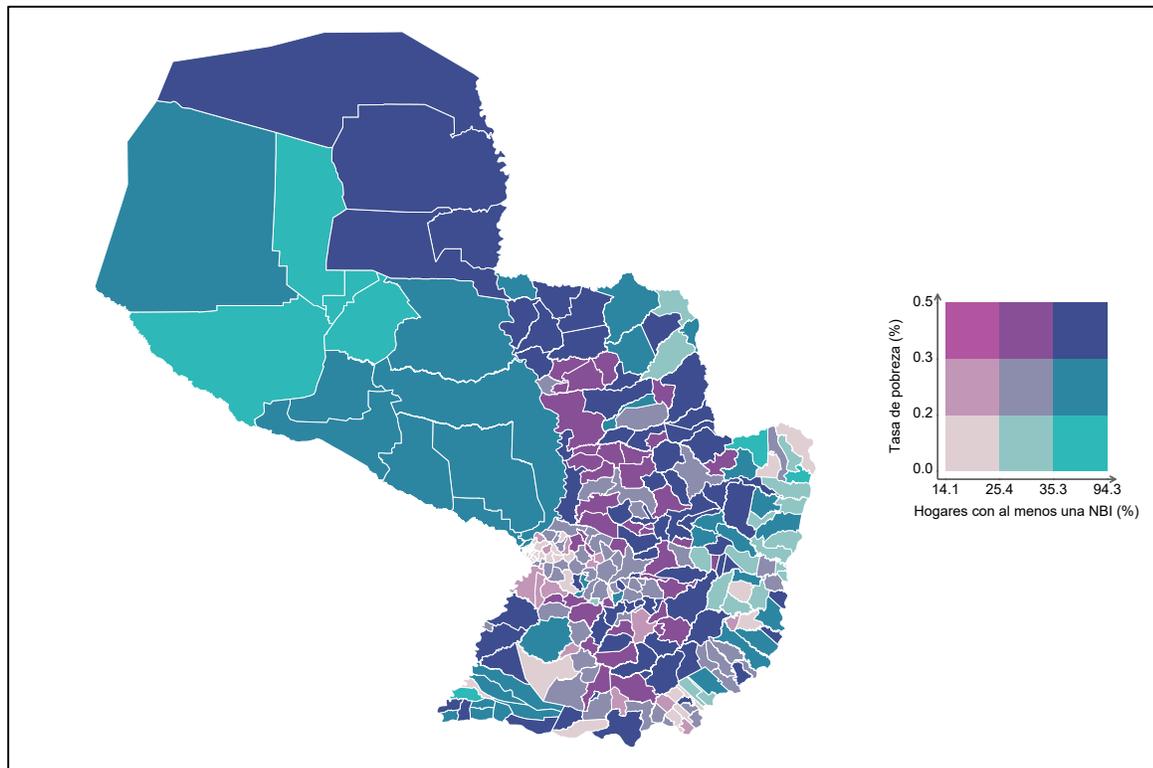
Para examinar los factores estructurales que subyacen a la distribución territorial de la pobreza, se realizó un análisis comparativo entre la pobreza monetaria estimada y las Necesidades Básicas Insatisfechas (NBI) recientemente publicadas por el INE (2025)¹¹. El índice de NBI contempla cuatro dimensiones estructurales: acceso a educación, calidad de vivienda, capacidad de subsistencia e infraestructura sanitaria. A nivel nacional, el 28,3% de los hogares presenta al menos una necesidad básica insatisfecha.

La Figura 7-3 revela patrones territoriales diferenciados en la relación entre privación monetaria y no monetaria. En la región del Chaco, particularmente en su sector noreste, se

¹¹ INE (2025). Necesidades Básicas Insatisfechas (NBI) Paraguay: Actualidad y su evolución en 4 décadas. Fernando de la Mora

observa una amplia superposición entre altas tasas de pobreza monetaria y elevados porcentajes de hogares con NBI, evidenciando un patrón de privación multidimensional estructural. Esta convergencia sugiere que los factores estructurales —acceso limitado a servicios básicos, infraestructura deficiente y oportunidades económicas restringidas— están altamente correlacionados con la pobreza monetaria en estos territorios.

Figura 7-3– Mapa de incidencia de la pobreza y necesidades básicas insatisfechas



Nota: Las estimaciones para Alto Paraguay y Boquerón son sintéticas (departamentos no incluidos en muestra EPHC 2023).

En contraste, la región Oriental presenta una variabilidad considerable en esta relación. Algunos distritos muestran altas tasas de pobreza monetaria con NBI relativamente menores, sugiriendo problemas de generación de ingresos en contextos con mejor infraestructura básica. Otros territorios evidencian el patrón inverso, indicando la persistencia de deficiencias estructurales en áreas con mejores condiciones económicas promedio.

Esta heterogeneidad territorial en la relación pobreza-NBI confirma que los factores estructurales operan de manera diferenciada según el contexto geográfico, validando la necesidad

de políticas territorialmente diferenciadas que aborden tanto las limitaciones de ingresos como las deficiencias en infraestructura y servicios básicos.

7.3 Otras posibles extensiones al mapa de pobreza

Los mapas de pobreza pueden combinarse con información geoespacial de diversa índole para potenciar su uso en políticas sociales. Por ejemplo, se pueden combinar con otras medidas de bienestar como mapas de pobreza multidimensional y mapas de necesidades básicas insatisfechas (NBI), de manera a tener un panorama más amplio de la pobreza en todas sus formas. También pueden complementarse con mapas de vulnerabilidad a la pobreza para entender mejor la distribución geográfica de los actualmente pobres y de quienes podrían caer en pobreza ante un choque adverso (Barriga- Cabanillas et al., 2024).

Los mapas de pobreza también pueden combinarse con mapas de riesgos de diversos eventos climáticos, como inundaciones, calor extremo y sequías, facilitando la identificación de las poblaciones más expuestas a riesgos climáticos. Asimismo, debido al rol fundamental de los servicios públicos en la lucha contra la pobreza, el mapeo de la relación entre la disponibilidad de estos servicios y la pobreza monetaria puede facilitar el diseño de mejores políticas para expandir infraestructura esencial.

8. Conclusiones

En este documento se presentan las estimaciones de tasas de pobreza para 263 distritos de Paraguay, y a nivel barrial para los 67 barrios de la capital del país aplicando el Mejor Estimador Empírico del Censo (EB) del método de áreas pequeñas, considerando que 1 barrio no tiene población. La estimación se basó en los datos de la Encuesta Permanente de Hogares Continua 2023 para luego imputar ingresos en el Censo de Población y Vivienda de 2022, para luego aplicar las líneas de pobreza oficial, y así obtener las tasas de pobreza estimadas en el censo. El análisis se complementa con un estudio territorial que examina los patrones de distribución espacial y su relación con factores estructurales. Es importante señalar que, los resultados de dicho análisis constituyen estadísticas experimentales que complementan las mediciones oficiales de pobreza y proporcionan una herramienta adicional para el diseño de políticas territorialmente diferenciadas. Entre los resultados más notables se encuentra lo siguiente:

- La incidencia de pobreza varía ampliamente a nivel geográfico, tanto entre departamentos, como entre distritos y barrios.
- El departamento con mayor tasa de pobreza en el 2022 es Alto Paraguay, con una tasa del 38,7%, en contraste, el departamento con menor tasa de pobreza es Boquerón, con un 8,3%.
- La tasa de pobreza a nivel de distrito muestra una variabilidad considerable: la tasa de incidencia de pobreza es mayor en el distrito de Paso Barreto, con una tasa del 53,3%, menor en el distrito de Fernando de la Mora, con un 3,7% de pobreza.
- La incidencia de pobreza en la capital del país es del 7,3%. Sin embargo, la tasa de pobreza es altamente heterogénea entre barrios, siendo el Barrio Banco San Miguel el de mayor incidencia con una tasa del 42,5%, en tanto que, en el otro extremo, el barrio con menor tasa de pobreza es el barrio Mburicao con 0,03%.
- El análisis territorial revela que la concentración absoluta de población pobre se localiza principalmente en centros urbanos y periurbanos, planteando desafíos diferenciados que requieren estrategias de focalización específicas según la combinación de incidencia de pobreza y tamaño poblacional.
- La comparación con indicadores de necesidades básicas insatisfechas confirma la influencia de factores estructurales en la distribución territorial de la pobreza, con patrones diferenciados entre la región del Chaco (convergencia crítica entre pobreza monetaria y carencias estructurales) y la región Oriental (heterogeneidad en esta relación).

Los hallazgos presentados en este estudio resaltan la necesidad de programas de reducción de la pobreza adaptados a las realidades locales del país. Las marcadas variaciones en los niveles de pobreza tanto dentro y entre departamentos y distritos enfatizan la importancia de la focalización subnacional en la oferta de servicios públicos. Las estimaciones de pobreza presentadas ofrecen usos relevantes para las políticas públicas: facilitan la identificación de tipologías territoriales para la focalización diferenciada de programas sociales, pueden combinarse con otros datos espaciales para explicar la persistencia de la pobreza en ciertas áreas geográficas, y permiten la priorización geográfica de intervenciones que buscan reducir la pobreza considerando tanto su intensidad como su magnitud absoluta. Finalmente, los mapas de pobreza constituyen herramientas fundamentales para orientar inversiones públicas sociales y avanzar hacia políticas de desarrollo territorial más efectivas y equitativas.

Bibliografía

- Banerjee, Abhijit V, Angus Deaton, Nora Lustig, Kenneth Rogoff, and Edward Hsu. 2006. "An Evaluation of World Bank Research, 1998-2005". In: Available at SSRN 2950327
- Barriga- Cabanillas, O., Bossuroy, T., Corral Rodas, P. A., Rodriguez Castelan, C., & Skoufias, E. 2024. *Sustaining Poverty Gains: A Vulnerability Map to Guide Social Policy* (No. 17193). IZA Discussion Papers. <https://www.iza.org/publications/dp/17193/sustaining-poverty-gains-a-vulnerability-map-to-guide-social-policy>
- Battese, G. E., Harter, R. M., and Fuller, W. A. 1988. "An Error Component Model for Predicting Counties' Cultivated Areas Using Survey Data and Satellite Data." *Journal of the American Statistical Association* 83 (401) 28-36. <http://www.jstor.org/stable/2288915>
- Corral P., Molina, I., y Nguyen, M. 2021. Pull your small area estimates up by the bootstraps. *Journal of Statistical Computation and Simulation*, 91(16), 3304-3357. <https://doi.org/10.1080/00949655.2021.1926460>
- Corral, P., Molina, I., Cojocarú, A., y Segovia, S. 2022. "Guidelines to Small Area Estimation for Poverty Mapping." Washington: World Bank. <https://documents1.worldbank.org/curated/en/099115306242236696/pdf/P1694340364c9803d0b7df097798bc42eac.pdf>
- Elbers, C., Lanjouw, J. O, y Lanjouw, P. 2003. "Micro-level welfare estimate." World Bank Policy Research Paper (2911). <http://documents.worldbank.org/curated/en/362131468739473297>
- Fay, Robert E. y Roger A. Herriot. 1979. "Estimates of Income for Small Places: An Application of James-Stein Procedures to Census Data". In: *Journal of the American Statistical Association* 74.366a, pp. 269–277. <https://doi.org/10.1080/01621459.1979.10482505>
- González-Manteiga, W., Lombardía, M. J., Molina, I., Morales, D. y Santamaría, L. 2008. "Bootstrap means error squared of a small EBLUP area." *Journal of Statistical Computation and Simulation*. 78(5), 443-462.
- INE. 2024. Principales Resultados. Pobreza Monetaria y Distribución del Ingreso EPHC 2023. Instituto Nacional de Estadísticas: Fernando de la Mora. URL: https://www.ine.gov.py/Publicaciones/Biblioteca/documento/246/Pobreza%20Monetaria_%20EPHC%202023_INE..pdf
- INE (2025). Necesidades Básicas Insatisfechas (NBI) Paraguay: Actualidad y su evolución en 4 décadas. Fernando de la Mora. <https://www.ine.gov.py/resumen/262/necesidades-basicas-insatisfechas-nbi-paraguay-2022>
- Lanjouw, P. y Yoshida, N. 2021. Poverty Monitoring Under Acute Data Constraints: A Role for Imputation Methods. *LIS Newsletter*. url: <https://www.lisdatacenter.org/newsletter/nl-2021-19-im-2/#>
- Molina, I. y Rao, J. 2010. "Small Area Estimation of Poverty Indicators." *Canadian Journal of Statistics* 38 (3) 369-385.

- Ministerio de Asuntos Exteriores, Unión Europea y Cooperación (2024). *Ficha País - Paraguay*. url: https://www.exteriores.gob.es/documents/fichaspais/paraguay_ficha%20pais.pdf
- Newhouse, D., Ramakrishnan, A., Swartz, T., Merfeld, J., y Lahiri, P. (2022). Small area estimation of monetary poverty in Mexico using satellite imagery and machine learning. *Oxford Bulletin of Economics and Statistics*.
- Nguyen, M. C., Corral, P., Azevedo, J. P., y Zhao, Q. 2018. "SAE: A Stata package for unit-level small area estimation." World Bank Policy Research Working Paper 8630.
- Robles, M. y Santander, H. 2004. Paraguay. Pobreza y Desigualdad de Ingresos a Nivel Distrital. DGEEC, Fernando de la Mora, Diciembre 2004. <https://www.ine.gov.py/Publicaciones/Biblioteca/documento/195/Pobreza%20y%20Desigualdad%20por%20distritos.pdf>
- Van der Weide, R. 2014. « GLS estimation and empirical by prediction for linear mixed models with heteroskedasticity and sampling weights: A background study for the povmap project. » World Bank Policy Research Working Paper 7028.
- Yoshida, N., Takamatsu, S., Yoshimura, K., Aron, D. V., Chen, X., Malgioglio, S., . . . Zhang, K. 2021. *The Concept and Empirical Evidence of SWIFT Methodology*. Washington, D.C. : The World Bank. url: <https://documents.worldbank.org/en/publication/documents-reports/documentdetail/099547109302235758/idu04a3a086c0a2da04853084b10e855253105f9>

Anexos

Anexo A: Lista de variables candidatas y criterios de selección

Variables candidatas	Criterios de selección inicial		
	Criterio 1: Similitud de preguntas	Criterio 2: Variación para explicar ingreso	Criterio 3: Comparación de Medias
<p>Area de residencia (urbana/rural)</p> <p>v_tipo_vivienda v_dormitorios v_tipo_pared v_tipo_piso v_tipo_techo v_agua_utilizada v_agua_medio v_tiene_celular v_tipo_desague v_cocina_con v_disposicion_basura v_lote_terreno_vivi v_tiene_comp_notebook v_tiene_tableta v_tiene_internet v_tiene_televisor v_tiene_lavarropa v_tiene_termocalefon v_tiene_aire v_tiene_microondas v_tiene_automovil v_tiene_motocicleta v_tipo_hogar p_edad p_relacion_parentesco psexo p_estado_civil p_sabe_leer p_anios_estudio p_grupo_empleo idioma Registro civil</p>	<p>EPH se refiere a la conexión domiciliaria. Censo se refiere a cualquier conexión.</p> <p>Censo solo refiere a pantalla plana, mientras que la EPHC a cualquier tipo de pantalla.</p> <p>Censo refiere a lavarropas automática, en tanto que la EPCH refiere a cualquier tipo de lavarropas.</p> <p>La interpretación y la capacitación fueron diferentes en el censo.</p> <p>Censo preguntó si habla alguno de los idiomas, en tanto que la EPHC refiere al idioma más hablado en el hogar</p> <p>No son estrictamente comparables</p>	<p>96% reporta 'casa, rancho'</p> <p>95% reporta 'si, tiene'</p>	<p>Proporciones ligeramente diferentes en el censo con respecto a la EPHC, pero que se mantuvo por su importancia para la estimación de la pobreza</p> <p>Diferente</p> <p>Diferente</p> <p>Diferente</p>
Total variables	27	25	22

Anexo B: Estadísticas de cobertura de área

Departamento	Areas	min	p25	p50	p75	max	media
Alto Paraguay	4	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%
Alto Paraná	22	0,6%	0,9%	1,0%	1,3%	2,2%	1,1%
Amambay	6	0,0%	0,0%	1,8%	2,2%	2,3%	1,3%
Asunción	67	0,0%	0,7%	0,9%	1,2%	2,3%	0,9%
Boquerón	4	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%
Caaguazú	22	0,0%	0,6%	0,8%	1,0%	1,4%	0,8%
Caazapá	11	1,8%	2,0%	2,2%	2,6%	2,7%	2,3%
Canindeyú	16	0,0%	0,7%	1,6%	1,9%	3,0%	1,4%
Central	19	0,4%	0,5%	0,5%	0,6%	0,8%	0,5%
Concepción	14	0,0%	0,0%	0,9%	1,2%	1,8%	0,7%
Cordillera	20	0,5%	0,9%	1,2%	1,3%	1,9%	1,2%
Guairá	18	0,0%	1,2%	1,4%	1,7%	2,2%	1,4%
Itapúa	30	0,0%	0,9%	1,2%	1,3%	2,2%	1,1%
Misiones	10	2,1%	2,5%	2,6%	2,8%	4,4%	2,7%
Paraguarí	18	0,0%	1,0%	1,2%	1,3%	1,4%	1,1%
Presidente Hayes	10	0,0%	1,7%	2,4%	3,1%	5,5%	2,4%
San Pedro	22	0,0%	1,0%	1,1%	1,4%	2,5%	1,1%
Ñeembucú	16	0,0%	1,9%	2,2%	3,3%	4,6%	2,5%

Anexo C: Pruebas de normalidad para transformación logarítmica

Estadísticas de asimetría y curtosis

Asimetría = 0,1443789

Curtosis = 3,57624

Prueba de asimetría y curtosis

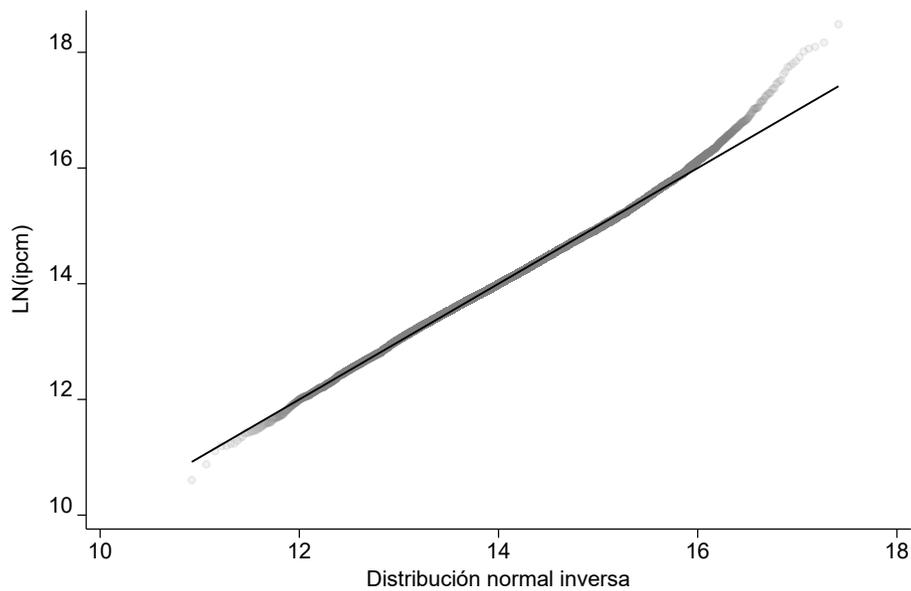
Variable	<i>N</i>	Pr(skewness)	Pr(kurtosis)	Adj. chi2(2)	Prob>chi2
lnipcm	17.027	0,0000	0,0000	175,82	0,0000

Prueba W de Shapiro-Wilk

Variable	<i>N</i>	W	V	z	Prob>z
lnipcm	17.027	0,99674	25,503	8,788	0,00000

Notas: Bajo la hipótesis nula de esta prueba la variable tiene una distribución normal. Valores *p* mayores a niveles de significancia convencionales (1%, 5%, o 10%), indican que los datos siguen una distribución normal.

Gráfico cuantil-cuantil con transformación logarítmica del ingreso



Anexo D: Pruebas de normalidad para transformación Box-Cox

Estadísticas de asimetría y curtosis

Asimetría = 0,00002

Curtosis = 3,493

Prueba de asimetría y curtosis

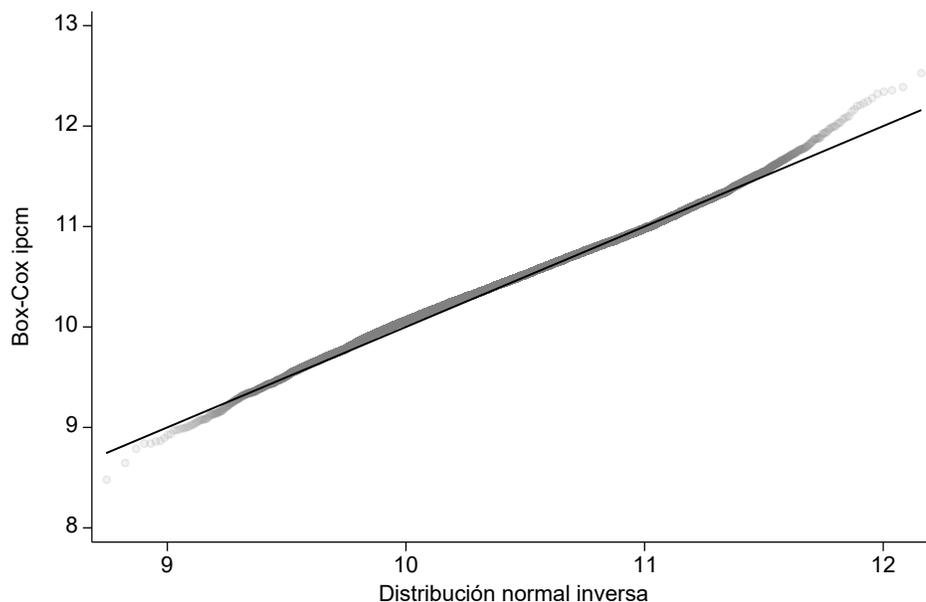
Variable	<i>N</i>	Pr(skewness)	Pr(kurtosis)	Adj. chi2(2)	Prob>chi2
lnipcm	17.027	0,9992	0,0000	99,69	0,0000

Prueba W de Shapiro-Wilk

Variable	<i>N</i>	W'	V'	z	Prob>z
bcipcm	17.027	0,99783	1,901	0,216	0,41449

Notas: Bajo la hipótesis nula de esta prueba la variable tiene una distribución normal. Valores *p* mayores a niveles de significancia convencionales (1%, 5%, o 10%), indican que los datos siguen una distribución normal.

Gráfico cuantil-cuantil con transformación Box-Cox del ingreso



Anexo E: Pruebas de normalidad para transformación logarítmica de sesgo cero

Estadísticas de asimetría y curtosis

Asimetría = 0,00007

Curtosis = 3,8226

Prueba de asimetría y curtosis

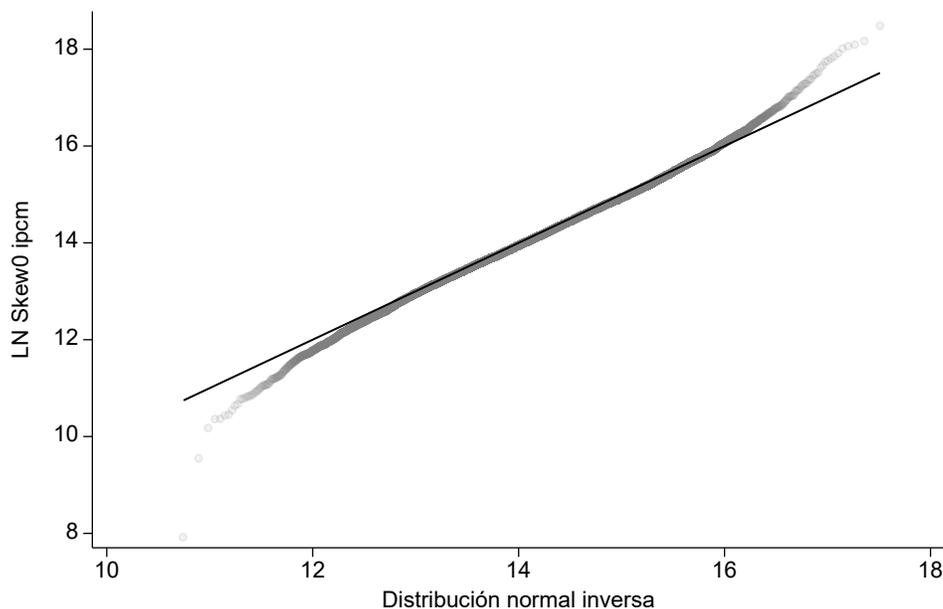
Variable	N	Pr(skewness)	Pr(kurtosis)	Adj. chi2(2)	Prob>chi2
lnipcm	17.027	0,9972	0,0000	214,16	0,0000

Prueba W de Shapiro-Wilk

Variable	N	W	V	z	Prob>z
lnipcm	17.027	0,99583	32,644	7,040	0,00000

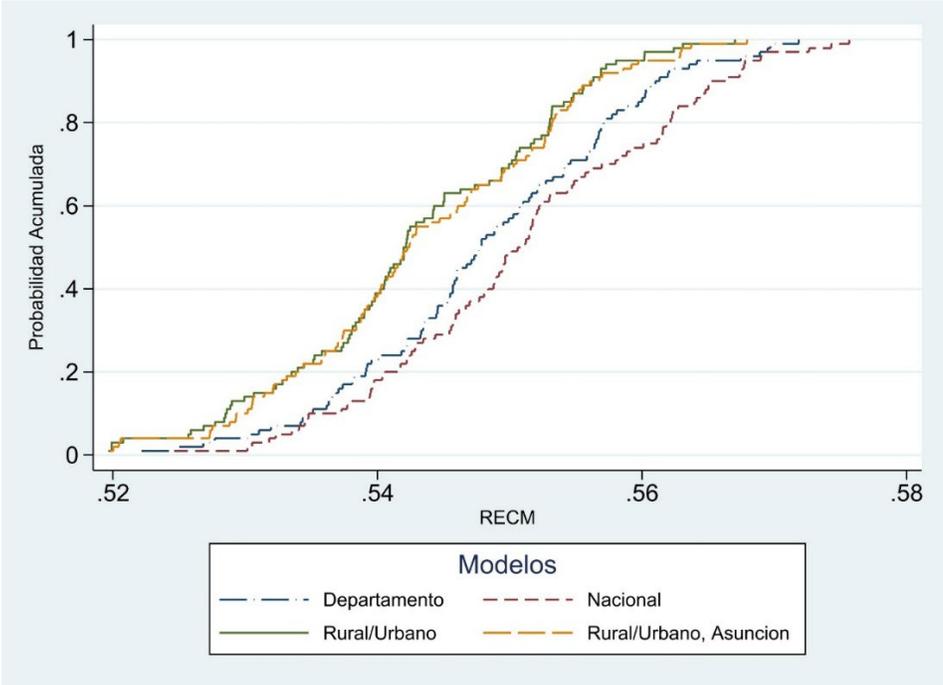
Notas: Bajo la hipótesis nula de esta prueba la variable tiene una distribución normal. Valores p mayores a niveles de significancia convencionales (1%, 5%, o 10%), indican que los datos siguen una distribución normal.

Gráfico cuantil-cuantil con transformación logarítmica de sesgo cero del ingreso



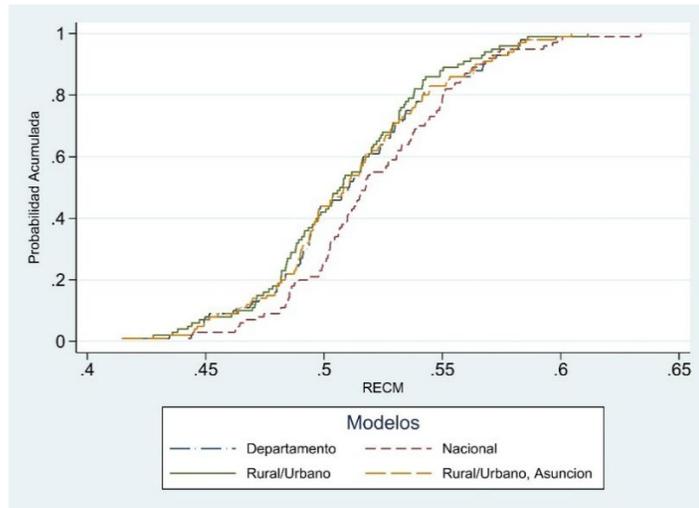
Anexo F: ECM de modelos alternativos

ECM Modelo Nacional

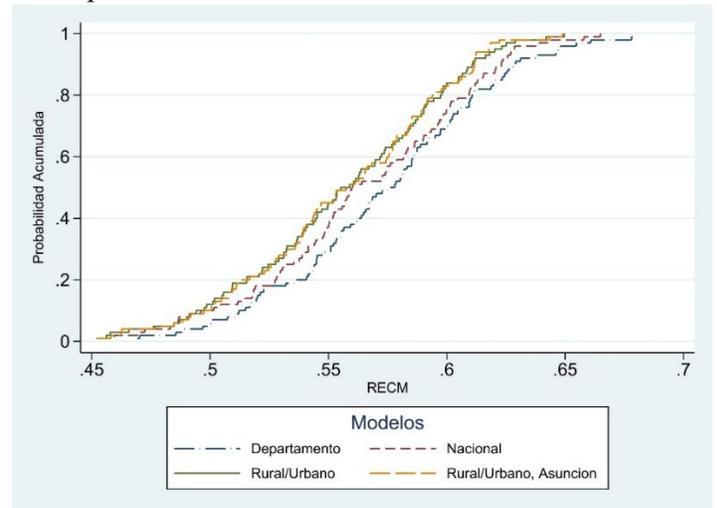


ECM Modelos Departamentales

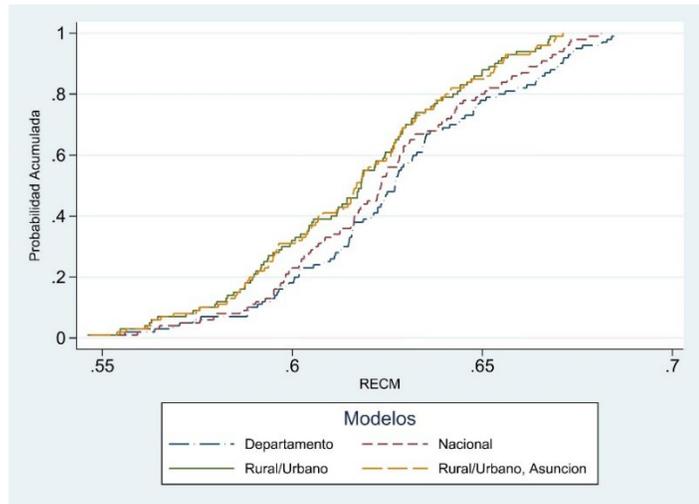
Asunción



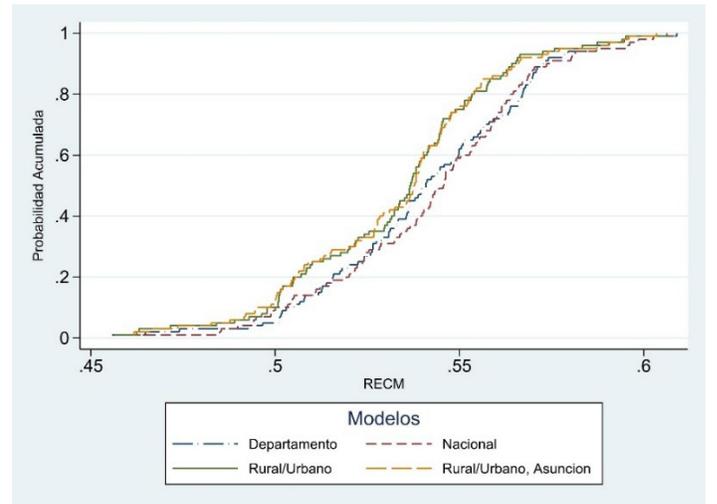
Concepción



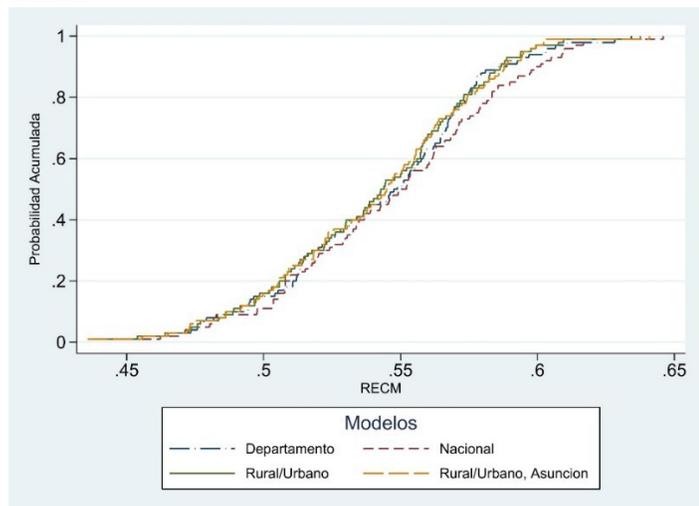
San Pedro



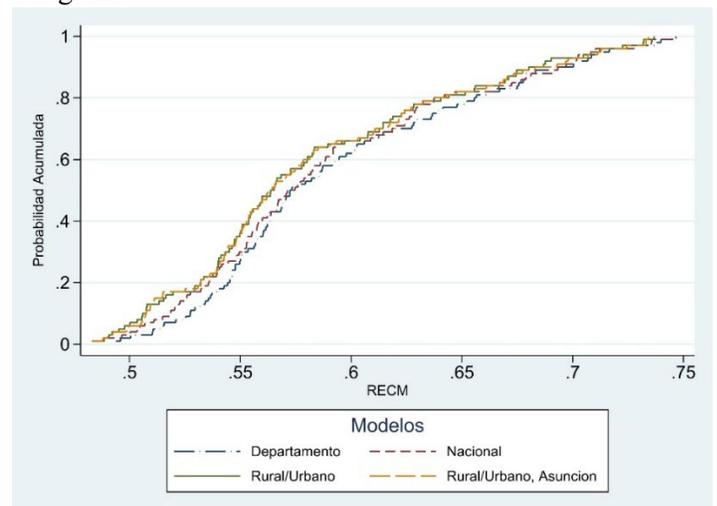
Cordillera



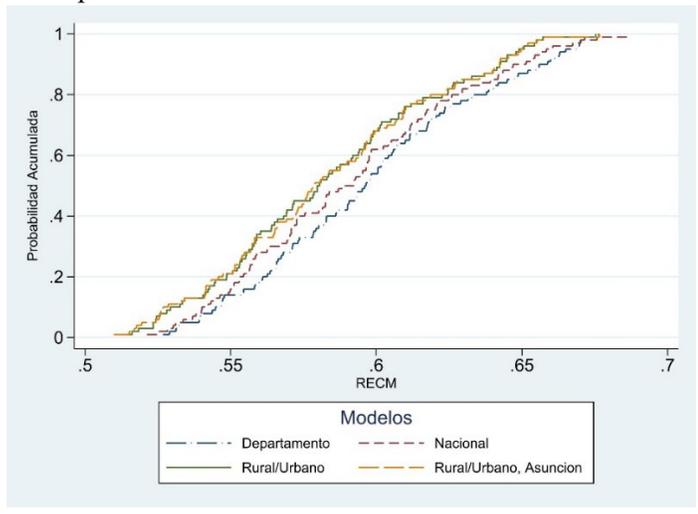
Guairá



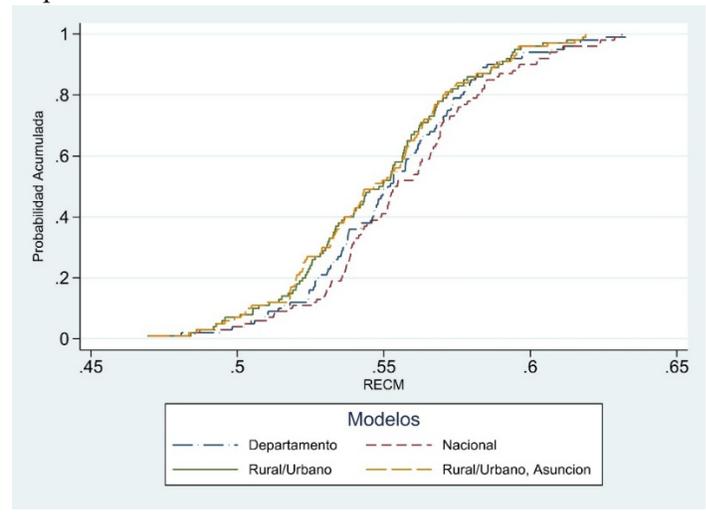
Caaguazú



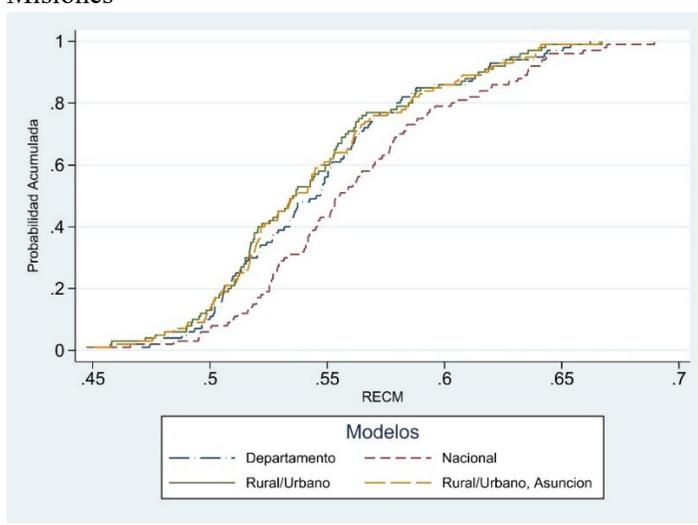
Caazapá



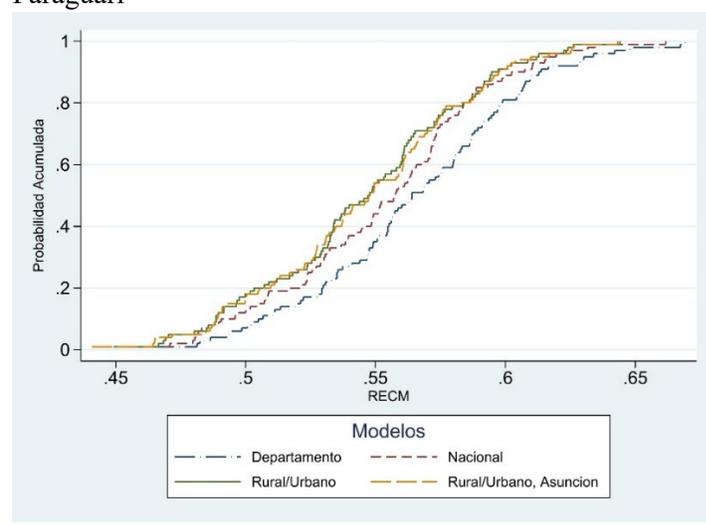
Itapúa



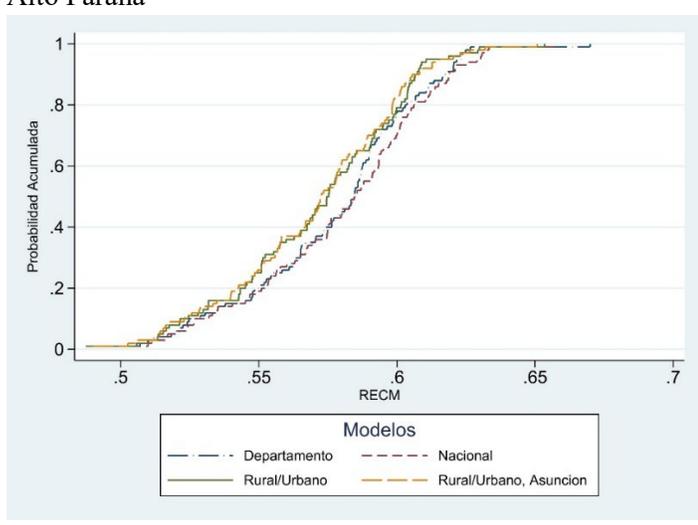
Misiones



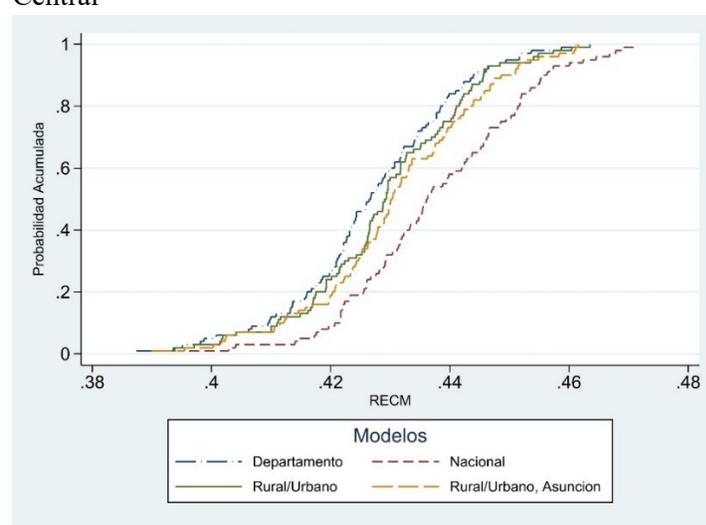
Paraguari



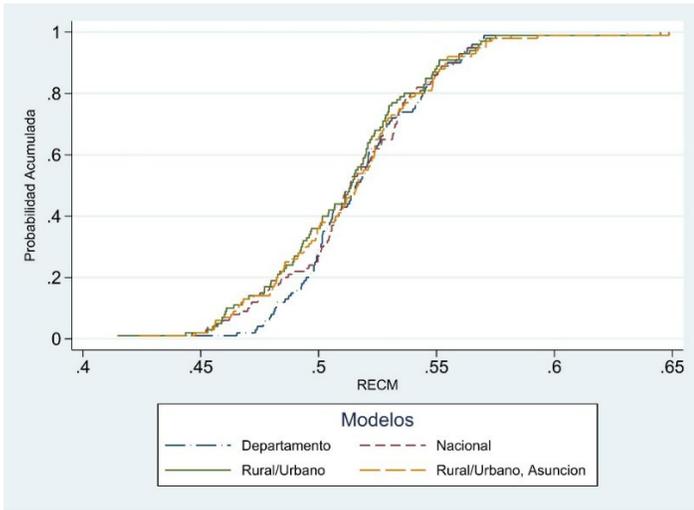
Alto Paraná



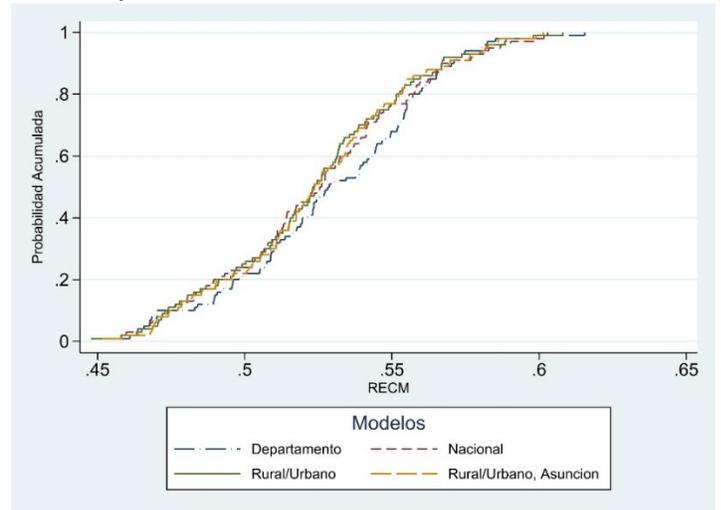
Central



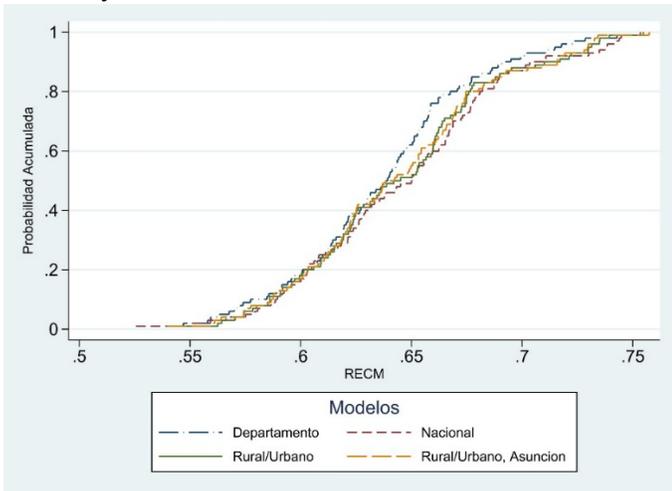
Ñeembucú



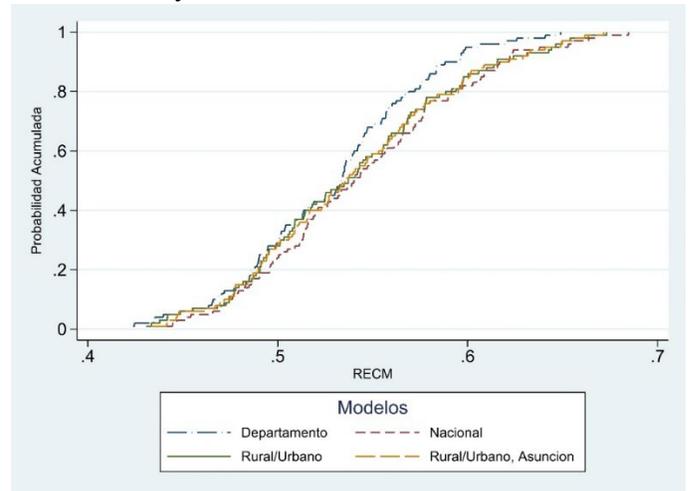
Amambay



Canindeyú



Presidente Hayes



Anexo G: Lista de variables y sus definiciones

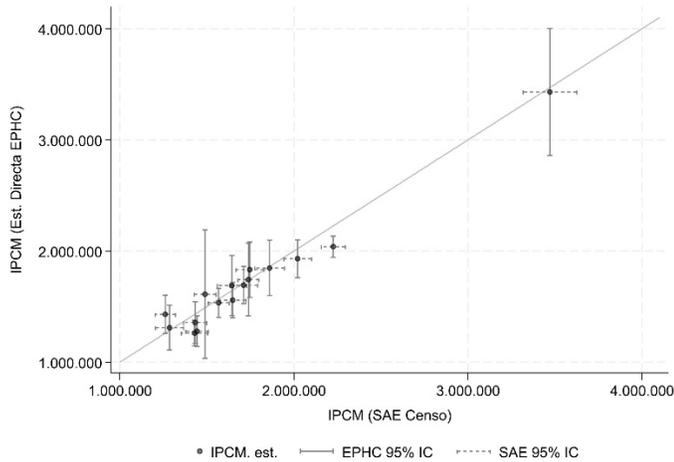
Nivel	Variable	Definición	Tipo de variable
Hogar, varía según área de residencia	area_rural	Toma el valor de 1 si el hogar vive en zona rural, de lo contrario 0	Dummy
	v_dormitorios_pc	Piezas utilizadas para dormir por persona en el hogar	Continua
	g_tipo_pared1	Toma el valor de 1 si el material de pared es de (1) bloque de cemento o (2) ladrillo, de lo contrario 0	Dummy
	g_tipo_techo1	Toma el valor de 1 si el material de techo es de (1) hormigón armado, loza o bovedilla, (2) teja, o (3) tablilla de madera, de lo contrario 0	Dummy
	g_agua_utilizada1	Toma el valor de 1 si el hogar tiene acceso a agua de (1) pozo artesiano, (2) ESSAP, (3) red o prestador privado, (4) pozo con bomba, o (5) SENASA o junta de saneamiento, de lo contrario 0	Dummy
	g_agua_medio1	Toma el valor de 1 si el hogar tiene acceso a agua de (1) cañería dentro de la vivienda, de lo contrario 0	Dummy
	g_tipo_combustible1	Toma el valor de 1 si el hogar cocina con (1) gas, (2) electricidad, o (3) no cocina, de lo contrario 0	Dummy
	g_disposicion_basural	Toma el valor de 1 si el hogar cuenta con recolección de basura pública/privada, de lo contrario 0	Dummy
	g_lote_terreno_vivi1	Toma el valor de 1 si el lote es (1) alquilado, (2) en condominio, (3) propio, o (4) lo están pagando en cuotas, de lo contrario 0	Dummy
	v_tiene_comp_notebook1	Toma el valor de 1 si el hogar tiene computadora o notebook, de lo contrario 0	Dummy
	v_tiene_aire1	Toma el valor de 1 si el hogar tiene acondicionador de aire, de lo contrario 0	Dummy
	v_tiene_microondas1	Toma el valor de 1 si el hogar tiene microondas, de lo contrario 0	Dummy
	v_tiene_automovill	Toma el valor de 1 si el hogar tiene automóvil, camión o camioneta, de lo contrario 0	Dummy
	v_tiene_motocicleta1	Toma el valor de 1 si el hogar tiene motocicleta, de lo contrario 0	Dummy
	g_unipersonal1	Toma el valor de 1 si el hogar tiene solo un miembro, de lo contrario 0	Dummy
	d_personas	Número de personas en el hogar	Continua
	j_edad	Edad del jefe de hogar	Continua
	d_porc_jovenes	Porcentaje de miembros del hogar menores de 25 años	Continua
	d_porc_mayores	Porcentaje de miembros del hogar mayores de 65 años	Continua
	j_feminino1	Toma el valor de 1 si el jefe de familia es una mujer, de lo contrario 0	Dummy
d_porc_mujeres	Porcentaje de miembros del hogar que son mujeres	Continua	
g_j_estado_civill	Toma el valor de 1 si el estado civil del jefe de hogar es separado o divorciado, de lo contrario 0	Dummy	
j_sabe_leer1	Toma el valor de 1 si el jefe de hogar sabe leer y escribir, de lo contrario 0	Dummy	
j_anios_estudio	Años de educación del jefe de hogar	Continua	
g_j_grupo_empleo1	Toma el valor de 1 si jefe (1) tiene empleo remunerado o (2) es inactivo jubilado o pensionado, de lo contrario 0	Continua	
d_porc_remunerados	Porcentaje de miembros que (1) tienen empleo remunerado o (2) son inactivo jubilado o pensionado	Continua	
Departamento	dpto_g_tipo_piso1	Proporción de los hogares con pisos de (1) parquet, (2) porcelanato, mármol, granito, o (3) baldosa común, mosaico, cerámica en el área	Continua
	dpto_g_tipo_techo1	Proporción de los hogares con techos de (1) hormigón armado, loza o bovedilla, (2) teja, o (3) tablilla de madera en el área	Continua
	dpto_g_tipo_desague1	Proporción de los hogares con desagüe de (1) red de alcantarillado sanitario (cloaca) en el área	Continua
	dpto_g_tipo_combustible1	Proporción de los hogares que cocina con (1) gas, (2) electricidad, o (3) no cocina en el área	Continua
	dpto_g_disposicion_basural	Proporción de los hogares con basura (1) recolección pública/privada en el área	Continua
	dpto_v_tiene_internet1	Proporción de los hogares con conexión a internet en el área	Continua
	dpto_v_tiene_comp_notebook1	Proporción de los hogares con computadoras o notebooks en el área	Continua
	dpto_v_tiene_tableta1	Proporción de los hogares con tablet o similares en el área	Continua
	dpto_v_tiene_aire1	Proporción de los hogares con acondicionador de aire en el área	Continua

Nivel	Variable	Definición	Tipo de variable
Distrito/barrio	distbar_pob_dens	Densidad de población por área	Continua
	distbar_d_personas	Tamaño medio de los hogares en el área	Continua
	distbar_g_unipersonal1	Proporción de hogares con un miembro en el área	Continua
	distbar_v_dormitorios_pc	Promedio de dormitorios por persona en los hogares del área	Continua
	distbar_g_tipo_pared1	Proporción de los hogares con paredes de (1) bloque de cemento o (2) ladrillo en el área	Continua
	distbar_g_tipo_piso1	Proporción de los hogares con pisos de (1) parquet, (2) porcelanato, mármol, granito, o (3) baldosa común, mosaico, cerámica en el área	Continua
	distbar_g_tipo_techo1	Proporción de los hogares con techos de (1) hormigón armado, loza o bovedilla, (2) teja, o (3) tablilla de madera en el área	Continua
	distbar_g_tipo_desague1	Proporción de los hogares con desagüe de (1) red de alcantarillado sanitario (cloaca) en el área	Continua
	distbar_g_tipo_combustible1	Proporción de los hogares que cocina con (1) gas, (2) electricidad, o (3) no cocina en el área	Continua
	distbar_g_disposicion_basural1	Proporción de los hogares con basura (1) recolección pública/privada en el área	Continua
	distbar_v_tiene_internet1	Proporción de los hogares con conexión a internet en el área	Continua
	distbar_v_tiene_comp_notebook1	Proporción de los hogares con computadoras o notebooks en el área	Continua
	distbar_v_tiene_tablet1	Proporción de los hogares con tablet o similares en el área	Continua
	distbar_v_tiene_aire1	Proporción de los hogares con acondicionador de aire en el área	Continua
	distbar_v_tiene_microondas1	Proporción de los hogares con microondas en el área	Continua
	distbar_v_tiene_automovil1	Proporción de los hogares con automóvil, camión o camioneta en el área	Continua
	distbar_v_tiene_motocicleta1	Proporción de los hogares con motocicleta en el área	Continua
	distbar_v_tiene_televisor1	Proporción de los hogares con televisor de pantalla plana en el área	Continua
	distbar_j_sabe_leer	Proporción de los jefes que saben leer y escribir en el área	Continua
	distbar_j_anios_estudio	Años promedio de escolaridad del jefe de hogar en el área	Continua
	distbar_j_feminino1	Proporción de mujeres jefas de hogar en el área	Continua
	distbar_zona_rural	Proporción de hogares en zonas rurales en el área	Continua
	distbar_perd_hogar	Proporción de hogares con información faltante o no respondida sobre las características de la vivienda en el área	Continua
	distbar_perd_activo	Proporción de hogares con información faltante o no respondida respecto de los activos en el área	Continua
	distbar_porc_jovenes	Porcentaje medio de miembros del hogar menores de 25 años en el área	Continua
	distbar_porc_adulto	Porcentaje medio de miembros del hogar de 25 a 65 años de edad en el área	Continua
	distbar_porc_remunerados	Porcentaje medio de miembros del hogar que (1) tienen empleo remunerado o (2) son inactivo jubilado o pensionado en el área	Continua

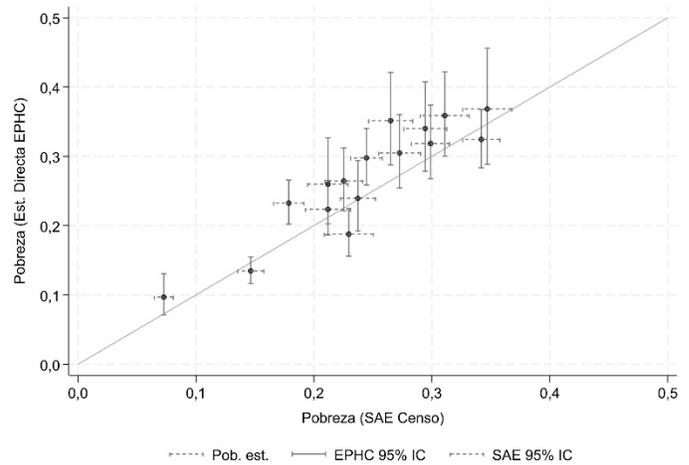
Anexo H: Diagramas de dispersión

Los diagramas de dispersión muestran una alta concordancia entre las clasificaciones departamentales obtenidas de las estimaciones del Censo SAE y las estimaciones directas del EPHC. La línea de 45° muestra igualdad. Los intervalos de confianza del 95% de casi todas las estimaciones cruzan esta igualdad, lo que sugiere una alta concordancia entre ellas.

a) Diagrama de dispersión del Ingreso estimado del Censo SAE y IPCM directo del EPHC



b) Diagrama de dispersión de la tasa de pobreza estimada del Censo SAE y la tasa de pobreza observada de la EPHC



Anexo I: Comparación de ratios a nivel departamental

Ratio de la media en la encuesta (EPHC 2023) y la media en el Censo 2022 para variables incluidas en el modelo SAE.

Variable común a la encuesta y al censo (sin incluir los promedios por corregimiento, distrito, o provincia)	Asunción	Concepción	San Pedro	Cordillera	Guairá	Caaguazú	Caazapá	Itapúa	Misiones	Paraguarí	Alto Paraná	Central	Ñeembucú	Amambay	Canindeyú	Presidente Hayes
g_tipo_pared1	0.988	1.010	0.928	0.988	0.982	0.974	0.875	1.031	1.006	0.996	1.014	0.987	0.982	1.032	0.959	0.973
g_tipo_techo1	0.960	1.094	0.907	0.951	0.950	0.768	0.829	0.952	0.967	0.951	1.089	0.942	0.894	1.035	0.970	0.822
g_agua_utilizada1	0.992	1.434	1.282	0.981	0.968	1.332	0.986	1.013	1.118	1.017	1.436	0.969	1.033	1.706	1.568	0.985
g_tipo_combustible1	0.985	1.017	0.738	0.879	0.907	0.873	0.820	0.976	0.940	0.830	0.967	0.985	0.977	0.982	0.845	0.969
g_disposicion_basura1	0.980	1.091	0.756	0.854	0.998	1.015	0.964	1.165	1.007	0.867	1.030	0.933	1.105	1.020	0.989	0.908
j_feminino1	0.978	1.051	1.215	1.086	1.150	1.138	1.253	1.133	1.061	1.115	1.030	1.017	1.110	1.167	1.105	0.996
j_sabe_leer1	1.007	1.070	1.054	1.036	1.031	1.070	1.045	1.049	1.049	1.055	1.033	1.011	1.042	1.068	1.067	1.047
v_tiene_comp_notebook1	1.239	1.807	0.905	1.007	1.197	1.232	0.942	1.389	1.244	1.055	1.226	1.113	1.235	1.067	1.112	1.069
v_tiene_aire1	1.008	1.034	0.729	0.861	0.939	0.907	0.810	1.095	1.007	0.942	1.087	0.994	0.999	0.789	0.920	1.021
v_tiene_microondas1	1.076	1.561	0.860	0.912	1.065	1.275	1.262	1.444	1.278	1.003	1.137	1.101	1.148	1.160	1.134	0.953
v_tiene_automovil1	1.146	1.141	0.799	0.992	1.075	1.186	0.923	1.115	1.055	1.058	1.188	1.096	1.022	1.242	1.077	1.057
ln_d_personas	0.951	1.002	1.005	0.970	0.950	0.992	0.954	0.963	0.978	0.936	0.979	0.967	1.027	1.031	1.013	0.937
d_porc_jovenes	0.974	1.009	1.037	1.009	0.952	1.007	1.011	1.003	1.050	0.930	0.967	1.027	1.149	1.055	1.015	0.972
d_porc_mayores	0.739	0.831	0.929	0.788	0.977	0.798	0.914	0.838	0.818	1.031	0.873	0.785	0.840	0.771	0.876	0.868
d_porc_remunerados	1.137	1.184	1.178	1.146	1.185	1.208	1.181	1.147	1.146	1.225	1.178	1.112	1.121	1.144	1.142	1.188
rur_v_tiene_automovil1		0.758	0.810	1.006	1.080	1.124	0.854	0.884	0.940	1.024	0.609	0.334	0.880	0.673	0.954	0.760
rur_v_tiene_motocicleta1		0.760	0.928	1.075	0.939	0.937	0.902	0.863	0.899	1.058	0.663	0.383	0.837	0.514	1.012	0.612
rur_ln_v_dormitorios_pc		24.955	0.819	0.955	1.108	0.826	1.375	0.850	0.756	1.527	3.537	0.051	0.463	-6.548	0.435	-3.669

Fuente: EPHC 2023 y Censo 2022



Naciones Unidas y Centeno,
Fernando de la Mora, Zona Norte



+595 986 800 506



info@ine.gov.py

