



PERÚ

Ministerio de Desarrollo
e Inclusión Social



INFORME DE EVALUACIÓN



Metodología para la priorización del servicio alimentario
del Programa Nacional de Alimentación Escolar Qali Warma

INFORME DE EVALUACIÓN

Título:

Metodología para la priorización del servicio alimentario del Programa Nacional de Alimentación Escolar Qali Warma

Elaboración:

Asociación Benéfica Prisma y Macroconsult S.A. Lima, Perú.

Año de elaboración:

2019

Eje de política:

Nutrición infantil y Desarrollo infantil temprano

Programas Sociales e Instrumentos de Política Social:

Programa Nacional de Alimentación Escolar Qali Warma

Resumen ejecutivo:

El estudio tiene como objetivo elaborar y validar los índices de priorización para el servicio alimentario del Programa considerando la situación socioeconómica y la situación nutricional de los estudiantes en el nivel de educación primaria. Como parte del estudio, se realizó el agrupamiento de las escuelas estableciendo tres grupos de interés: alta, media y baja necesidad, para el índice socioeconómico (utilizando quintiles de riqueza y análisis por conglomerados) y alta, media y baja prevalencia de sobrepeso, para el índice de la situación nutricional (considerando umbrales establecidos por la Organización Mundial de la Salud - OMS), dando origen a una matriz de 9 posibles categorías. La clasificación de las escuelas de acuerdo a su condición socioeconómica y situación nutricional posibilitaría distinguir el servicio alimentario que el PNAEQW brinda a cada escuela, así como las modalidades de atención: desayunos y almuerzos. El estudio también propone considerar tres modalidades de atención: (i) desayunos, (ii) almuerzos y (iii) refrigerios, en base al tipo de clasificación de cada escuela de acuerdo a sus criterios socioeconómicos y nutricionales. Esta nueva propuesta de diferenciación del servicio y la clasificación de las escuelas mediante criterios de priorización, permitiría además reducir el presupuesto del PNAEQW, al reducir la cantidad de alumnos que recibe desayuno y almuerzo e incrementar la cantidad de alumnos que recibe desayuno o refrigerio. Finalmente, el estudio propone un protocolo de actualización de los índices de priorización considerando una frecuencia quinquenal.

Supervisión:

Programa Mundial de Alimentos (PMA), con los aportes de la DGSE del MIDIS a los términos de referencia y al informe final del estudio.

Reservados algunos derechos:

Elaborado por la Asociación Benéfica Prisma y Macroconsult S.A. por encargo de la Oficina del PMA en el Perú. Las opiniones, interpretaciones y conclusiones aquí expresadas no son necesariamente reflejo de la opinión del PMA o del MIDIS. Nada de lo establecido en este documento constituirá o se considerará una limitación o renuncia a los privilegios del MIDIS, los cuales se reservan específicamente en su totalidad.

Cita de fuente:

El documento debe citarse de la siguiente manera: "PMA (2019). Metodología para la priorización del servicio alimentario del Programa Nacional de Alimentación Escolar Qali Warma. Informe de evaluación. Elaborado por la Asociación Benéfica Prisma y Macroconsult S.A. Lima, Perú".

PROGRAMA DE LAS NACIONES UNIDAS PARA EL DESARROLLO (PNUD)

OFICINA DEL PROGRAMA MUNDIAL DE ALIMENTOS EN EL PERÚ (PMA)

PROYECTO PER18RFP062 :

**“METODOLOGÍA PARA LA PRIORIZACIÓN DEL SERVICIO ALIMENTARIO DEL
PROGRAMA NACIONAL DE ALIMENTACIÓN ESCOLAR QALI WARMA”**

CONSULTOR: ASOCIACIÓN BENÉFICA PRISMA

ENTREGABLE 4 DE 4 – ÍNDICES DE PRIORIZACIÓN

VERSIÓN FINAL

Lima, 27 de junio de 2019

Metodología para la priorización del servicio alimentario del Programa Nacional de Alimentación Escolar Qali Warma

INFORME FINAL

Junio de 2019

Este documento fue preparado por la Asociación Benéfica Prisma y Macroconsult S.A. para la Oficina del Programa Mundial de Alimentos en el Perú.

Metodología para la priorización del servicio alimentario del Programa Nacional de Alimentación Escolar Qali Warma

INFORME FINAL

Junio de 2019

Las opiniones de los autores consignados en este documento no reflejan, necesariamente, los puntos de vista de la Oficina del Programa Mundial de Alimentos en el Perú.

LISTA DE ACRÓNIMOS

CAE	Comité de Alimentación Escolar
CC	Comité de Compras
CENAN	Centro Nacional de Alimentación y Nutrición
Censo 2017	Censos Nacionales 2017: XXII de Población, VII de Vivienda y III de Comunidades Indígenas
CEPAL	Comisión Económica para América Latina y el Caribe
DGSE	Dirección General de Seguimiento y Evaluación del MIDIS
DNI	Documento Nacional de Identidad
ECMP	Error cuadrático medio de predicción
ENAHO	Encuesta Nacional de Hogares
ENDES	Encuesta Demográfica y de Salud Familiar
IMC	Índice de Masa Corporal
INEI	Instituto Nacional de Estadística e Informática
MEF	Ministerio de Economía y Finanzas
MIDIS	Ministerio de Desarrollo e Inclusión Social
MINEDU	Ministerio de Educación
MINSA	Ministerio de Salud
MPI	Índice de Pobreza Multidimensional
NBI	Necesidades Básicas Insatisfechas
OMS	Organización Mundial de la Salud
OPHI	Oxford Poverty and Human Development Initiative
PCA	Análisis de Componentes Principales
PET	Población en Edad de Trabajar
PMA	Oficina del Programa Mundial de Alimentos en el Perú
PNAEQW	Programa Nacional de Alimentación Escolar Qali Warma
PSE	Plan de Salud Escolar
SIAGIE	Sistema de Información de Apoyo a la Gestión de la Institución Educativa
SIS	Seguro Integral de Salud
SISFOH	Sistema de Focalización de Hogares

CONTENIDO

LISTA DE ACRÓNIMOS	4
ÍNDICE DE GRÁFICOS	6
ÍNDICE DE MAPAS	6
ÍNDICE DE TABLAS	7
RESUMEN EJECUTIVO.....	8
1. INTRODUCCIÓN	13
2. CONTEXTO DEL ESTUDIO.....	15
3. OBJETIVOS DEL ESTUDIO.....	18
4. RESUMEN METODOLÓGICO.....	19
5. RESULTADOS DE LA PRIORIZACIÓN.....	22
6. MAPAS DE PRIORIZACIÓN 2018	28
7. LIMITACIONES DEL ESTUDIO	32
8. PROTOCOLO DE ACTUALIZACIÓN DEL ÍNDICE.....	33
9. CONCLUSIONES	45
10. BIBLIOGRAFÍA.....	48
11. ANEXOS	50
11.1. ANEXO METODOLÓGICO.....	51
11.2. VALIDACIÓN DE LOS ÍNDICES DE PRIORIZACIÓN	87
11.3. ARCHIVOS DIGITALES	89

ÍNDICE DE GRÁFICOS

Gráfico 1. Prevalencia de la desnutrición infantil crónica, 2008-2018	16
Gráfico 2. Evolución de la prevalencia del sobrepeso y obesidad infantil en niños de 5 a 9 años, 2007-2014	17
Gráfico 3. Evolución de la prevalencia del sobrepeso y obesidad infantil en niños de 10 a 19 años, 2007-2014.....	17
Gráfico 4. Estado nutricional de los alumnos de primaria según el PSE, 2015-2016.....	20
Gráfico 5. Metodología para la priorización del PNAEQW.....	21
Gráfico 6. Modalidades de atención sugeridas por el PMA.....	23
Gráfico 7. Distribución de las modalidades de atención a nivel escuelas.....	27
Gráfico 8. Actores relevantes para la actualización de los índices de priorización	40
Gráfico 9. Correlación entre el gasto real per cápita y la tasa de pobreza según NBI a nivel departamental.....	65
Gráfico 10. Correlación entre el gasto real per cápita y el puntaje del primer componente principal a nivel departamental	65
Gráfico 11. Distribución de los índices socioeconómicos observada e imputada	70
Gráfico 12. Distribución de los indicadores socioeconómicos.....	73
Gráfico 13. Estado nutricional de los alumnos de primaria según el PSE, 2015-2016.....	74
Gráfico 14. Distribución de los indicadores de sobrepeso y obesidad	75
Gráfico 15. Distribución de los indicadores de sobrepeso y obesidad	76
Gráfico 16. Distribución de las escuelas según método de cálculo	80
Gráfico 17. Pronóstico fuera de la muestra bajo cada modelo	84
Gráfico 18. Distribución de la tasa de alumnos con sobrepeso y obesidad.....	85
Gráfico 19. Correlación entre índices nutricionales.....	88

ÍNDICE DE MAPAS

Mapa 1. Distribución de escuelas según índice socioeconómico	29
Mapa 2. Distribución de escuelas según índice nutricional	30
Mapa 3. Mapa de priorización del PNAEQW 2018	31
Mapa 4. Distribución de las escuelas según necesidades de imputación.....	68

ÍNDICE DE TABLAS

Tabla 1. Distribución de escuelas y de alumnos según índices de priorización*	10
Tabla 2. Categorías sugeridas.....	11
Tabla 3. Distribución de escuelas y de alumnos según índices de priorización y categorías sugeridas*	11
Tabla 4. Categorías sugeridas.....	22
Tabla 5. Distribución de escuelas y de alumnos según índices de priorización*	24
Tabla 6. Distribución de escuelas y de alumnos según índices de priorización y categorías sugeridas*	25
Tabla 7. Clasificación actual del PNAEQW para la atención de escuelas*	26
Tabla 8. Comparación de las modalidades de atención.....	27
Tabla 9. Distribución de escuelas y de alumnos según índices de priorización*	46
Tabla 10. Resultados del emparejamiento de bases de datos: PNAEQW 2019, SIAGIE 2017, Censo 2017 y PSE	55
Tabla 11. Indicadores de NBI a nivel de alumnos	59
Tabla 12. Análisis de componentes principales y estadísticas descriptivas de indicadores	63
Tabla 13. Necesidades de imputación – índice socioeconómico	67
Tabla 14. Distribución de las distancias a las escuelas más cercanas	69
Tabla 15. Clasificación de las escuelas según índices socioeconómicos	72
Tabla 16. Clasificación de las escuelas según índices socioeconómicos – distribución de alumnos.....	72
Tabla 17. Clasificación de la información a nivel de escuelas	78
Tabla 18. Clasificación de las escuelas según métodos de cálculo del indicador nutricional	79
Tabla 19. Resultados de los modelos de estimación.....	83
Tabla 20. Distribución de las distancias a las escuelas más cercanas	85
Tabla 21. Clasificación de las escuelas según el índice nutricional	86
Tabla 22. Modelos de ajuste de los índices nutricionales.....	88

RESUMEN EJECUTIVO

El Programa Nacional de Alimentación Escolar Qali Warma (en adelante, **PNAEQW**) es un programa universal que brinda el servicio de alimentación escolar con complemento educativo a niños matriculados en instituciones públicas de inicial y primaria y, secundaria (en el caso de las poblaciones de la Amazonía Peruana). El objetivo del PNAEQW es asegurar la alimentación de los alumnos en la escuela durante todos los días del año escolar, de modo que incrementen su atención en las clases, aumenten su asistencia a la escuela y mejoren sus hábitos alimenticios.

El servicio que ofrece el programa se entrega de manera diferenciada según quintiles de pobreza. A los alumnos de las escuelas de los distritos más pobres del país (quintiles 1 y 2) se les entrega dos raciones de alimentos (desayuno y almuerzo); mientras que, a los alumnos de las escuelas de los distritos más ricos (quintiles 3, 4 y 5) se les entrega solo una ración (desayuno). A pesar de considerar la dimensión socioeconómica en la provisión del servicio, todas las escuelas dentro de un mismo distrito reciben el mismo servicio, independientemente del nivel de vulnerabilidad particular de cada escuela. Aunado a lo anterior, se tiene que la provisión del servicio no toma en cuenta la severidad del principal problema que PNAEQW intenta resolver: la inadecuada ingesta de alimentos. En este contexto, surge la necesidad de priorizar el servicio alimentario que ofrece el PNAEQW, y en 2018 la Oficina del Programa Mundial de Alimentos en el Perú (en adelante, **PMA**) encarga la elaboración de una Nota Metodológica para la priorización del servicio alimentario del PNAEQW (ver Lavado y otros (2018)).

El objetivo del presente estudio es elaborar y validar los índices de priorización propuestos en dicha Nota Metodológica para la prestación del servicio alimentario del PNAEQW en el nivel de educación de primaria. Es decir, un índice que considere la situación socioeconómica de las escuelas y otro que considere la situación nutricional. Para alcanzar este objetivo se realizaron tres actividades: (1) armonización de las bases de datos; (2) construcción de los índices de priorización; y (3) clasificación de las escuelas en grupos de priorización. Primero, se armonizaron cuatro bases de datos: el padrón de escuelas atendidas por el PNAEQW en el año 2019, el padrón de alumnos del Sistema de Información de Apoyo a la Gestión de la Institución Educativa (SIAGIE) 2017, el padrón de población y vivienda de los Censos Nacionales del 2017, y el Plan de Salud Escolar (en adelante, **PSE**) 2015-2016. El universo de información viene dado por el padrón del PNAEQW, donde se registran 29 mil 486 escuelas públicas de nivel primario atendidas en la actualidad que benefician a cerca de 2.6 millones de alumnos. Como resultado del empate realizado, se cuenta con información socioeconómica de los Censos Nacionales 2017: XXII de Población, VII de Vivienda y III de Comunidades Indígenas (en adelante, **Censo 2017**) del 98% de escuelas (28,742 escuelas) y del 72% de alumnos (1,804,232 alumnos); y con información de talla y peso del 91% de escuelas (26,787 escuelas) y del 42% de los alumnos (1,038,754).

Con la base de datos anterior, se construyeron los índices de priorización del estudio. En ambos casos, el procedimiento fue el siguiente: primero, se construyeron los indicadores a nivel de alumnos y luego se agregaron a nivel de escuela. En el caso del índice socioeconómico, se implementaron dos metodologías para su construcción, una a partir de la tasa de pobreza según

el método de las Necesidades Básicas Insatisfechas (en adelante, **NBI**), y otra partir de un índice de riqueza construido a través de un ejercicio de Análisis de Componentes Principales (en adelante, **PCA**). En el caso del índice nutricional, se utilizó una de las sugerencias de la Nota Metodológica, el Índice de Masa Corporal (en adelante, **IMC**), a partir del cual se calculó la carga de malnutrición de cada escuela. En el Perú, esta última se concentra fundamentalmente en el sobrepeso y la obesidad por lo que ese fue el indicador seleccionado junto con el equipo técnico del PMA. De acuerdo con los resultados del PSE, la delgadez y la delgadez extrema –la otra cara de la malnutrición–, afectan al 2% de los estudiantes evaluados, por lo que esta condición no constituye un indicador relevante para efectos de la priorización. A continuación, se estimó la proporción de escuelas que requiere algún procedimiento de imputación. En el caso del índice socioeconómico, se requirió imputar información al 6% de escuelas que atiende el PNAEQW (1,934 escuelas) a partir del valor promedio de las escuelas vecinas más cercanas (según distancia geográfica); mientras que en el caso del índice nutricional fue necesario imputar al 32% de las escuelas a partir de una estrategia que, en una primera etapa, emplea técnicas de predicción de *machine learning* y, en una segunda etapa, emplea la misma estrategia de cercanía de escuelas vecinas que en el índice socioeconómico.

Con la data completa, se realizó el agrupamiento de las escuelas a partir cada índice. De manera *a priori*, se establecieron tres grupos de interés para el índice socioeconómico (alta, media y baja necesidad) y tres para el índice nutricional (alta, media y baja prevalencia de sobrepeso y obesidad). En el estudio se presentan tres ejercicios de agrupamiento para el caso del índice socioeconómico (dos ejercicios utilizando quintiles de riqueza y un análisis por conglomerados) y uno para el nutricional (considerando los umbrales establecidos por la Organización Mundial de la Salud (OMS) para identificar a la anemia como un problema de salud pública, así como la distribución de la carga de malnutrición de las escuelas; este criterio fue definido junto con el equipo técnico del PMA). Finalmente, para cada una de las opciones presentadas, se cruzó el índice socioeconómico con el nutricional dando origen, a una matriz donde se distinguen nueve posibles categorías.

En breve, en el estudio se presentan seis alternativas para priorizar a las escuelas, las cuales se diferencian por la forma en la que se construye el índice socioeconómico (NBI o PCA) y se realiza el agrupamiento de escuelas (quintiles de riqueza, quintiles de riqueza incorporando a las poblaciones especiales, o análisis por conglomerados). La decisión final sobre qué metodología emplear debería estar a cargo del equipo técnico del PNAEQW. No obstante, sugerimos utilizar aquella que utiliza el índice de riqueza construido a partir del PCA, toda vez que se encuentra más correlacionado con el gasto del hogar a comparación del índice basado en las NBI y, en ese sentido, puede ser reemplazado fácilmente por el Mapa de Pobreza que actualmente elabora el Instituto Nacional de Estadística e Informática (INEI). Asimismo, sugerimos agrupar las escuelas a partir de una estrategia de análisis por conglomerados (*k-medias*), debido a que esta técnica no solo reduce la discrecionalidad en el agrupamiento de escuelas a comparación de las otras alternativas, sino que ofrece un agrupamiento que maximiza la homogeneidad de las escuelas al interior de cada grupo y la heterogeneidad de estos al exterior. Tomando en cuenta este índice y el agrupamiento por análisis de conglomerados se tiene la siguiente distribución de escuelas:

Tabla 1. Distribución de escuelas y de alumnos según índices de priorización*

Clasificación según índice nutricional	Clasificación según índice socioeconómico			Total
	Necesidad alta	Necesidad media	Necesidad baja	
Cantidad y porcentaje de escuelas				
Sobrepeso/obesidad alta	1,641 6%	2,232 8%	3,101 11%	6,974 24%
Sobrepeso/obesidad media	10,192 35%	3,816 13%	258 1%	14,266 48%
Sobrepeso/obesidad baja	6,887 23%	1,237 4%	122 0%	8,246 28%
Total	18,720 63%	7,285 25%	3,481 12%	29,486 100%
Cantidad y porcentaje de alumnos				
Sobrepeso/obesidad alta	67,699 3%	319,742 12%	1,160,855 45%	1,548,296 60%
Sobrepeso/obesidad media	383,265 15%	363,162 14%	81,811 3%	828,238 32%
Sobrepeso/obesidad baja	172,723 7%	33,590 1%	7,223 0%	213,536 8%
Total	623,687 24%	716,494 28%	1,249,889 48%	2,590,070 100%

*Nota: Los porcentajes están expresados respecto al total de escuelas y de alumnos atendidos por el PNAEQW.

Fuente: Base de datos armonizada. **Elaboración propia.**

La clasificación de las escuelas de acuerdo a su condición socioeconómica y nutricional permitirá diferenciar el servicio alimentario que el PNAEQW entrega a cada escuela o, por lo menos, las modalidades de atención. Actualmente, el programa atiende a partir de dos modalidades: desayunos y almuerzos. Para el presente estudio se consideraron tres modalidades de atención¹: (i) desayunos; (ii) desayunos + almuerzos; y (iii) refrigerios (ver **Tabla 2**). Los desayunos se entregarían a los alumnos de las escuelas clasificadas como de alta necesidad y que tienen alto sobrepeso/obesidad, y a las escuelas clasificadas como de media necesidad y sobrepeso/obesidad media y baja. Los desayunos + almuerzos se entregarían a los alumnos de las escuelas clasificadas como de alta necesidad y sobrepeso/obesidad media y baja. Finalmente, los refrigerios, se entregarían al resto de escuelas, constituyendo así la modalidad base.

¹ Estas modalidades de atención fueron sugeridas por el PMA sobre la base de la experiencia internacional y la asistencia técnica que el Programa Mundial de Alimentos brinda a los programas de alimentación escolar en el mundo. Estas modalidades no solo hacen referencia al tipo de menú que programa ofrecerá (desayunos, desayunos + almuerzos, o refrigerios), sino también a las características de estos (tipos de combinaciones, grupos de alimentos y proporciones), así como al componente educativo que ofrece PNAEQW.

Tabla 2. Categorías sugeridas

Categorías del índice nutricional	Categorías del índice socioeconómico		
	Necesidad alta	Necesidad media	Necesidad baja
Sobrepeso/obesidad alta	Desayuno	Refrigerio	Refrigerio
Sobrepeso/obesidad media	Desayuno + almuerzo	Desayuno	Refrigerio
Sobrepeso/obesidad baja	Desayuno + almuerzo	Desayuno	Refrigerio

Elaboración propia.

Dado que en su versión actual (a la fecha de cierre del presente informe) la propuesta de modalidades de atención es la misma para las escuelas sobrepeso/obesidad media y baja, el contenido de la **Tabla 1** puede resumirse en seis categorías tal como se presenta en la **Tabla 3**. La clasificación de las escuelas de acuerdo a criterios socioeconómicos y nutricionales y la consecuente diferenciación del servicio que se deriva de ella, permitiría, además, reducir el presupuesto del PNAEQW, toda vez que se reduciría la cantidad de alumnos que recibiría desayuno + almuerzo (la modalidad más costosa) e incrementaría la cantidad de alumnos que recibiría desayuno o refrigerio. No obstante, es importante volver a mencionar que estas nuevas modalidades de atención solo constituyen una sugerencia, y que deben ser validadas por el equipo técnico del PNAEQW.

Tabla 3. Distribución de escuelas y de alumnos según índices de priorización y categorías sugeridas*

Clasificación según índice nutricional	Clasificación según índice socioeconómico			Total
	Necesidad alta	Necesidad media	Necesidad baja	
Cantidad y porcentaje de escuelas				
Sobrepeso/obesidad alta	1,641	2,232	3,101	6,974
	6%	8%	11%	24%
Sobrepeso/obesidad media y baja	17,079	5,053	380	22,512
	58%	17%	1%	78%
Total	18,720	7,285	3,481	29,486
	63%	25%	12%	100%
Cantidad y porcentaje de alumnos				
Sobrepeso/obesidad alta	67,699	319,742	1,160,855	1,548,296
	3%	12%	45%	60%
Sobrepeso/obesidad media y baja	555,988	396,752	89,034	1,041,774
	22%	15%	3%	40%
Total	623,687	716,494	1,249,889	2,590,070
	24%	28%	48%	100%

*Nota: Los porcentajes están expresados respecto al total de escuelas y de alumnos atendidos por el PNAEQW.

Fuente: Base de datos armonizada. Elaboración propia.

Finalmente, como parte de este estudio se elaboró un protocolo de actualización de los índices de priorización. El protocolo fue elaborado considerando una frecuencia de actualización quinquenal y bajo el supuesto de que el Estado contará con un Mapa de Pobreza actualizado a partir del Censo Nacional y de que continuará realizando mediciones de talla y peso a todos los alumnos matriculados en instituciones públicas con cierta periodicidad. Caso contrario, no será posible actualizar ambos índices, al menos, con la frecuencia mencionada. La institución responsable de la actualización del índice debería ser el Ministerio de Desarrollo e Inclusión Social (MIDIS), quien necesitará contar con el apoyo del INEI para la armonización de la información requerida.

1. INTRODUCCIÓN

El Programa Nacional de Alimentación Escolar Qali Warma (en adelante, **PNAEQW**) es el programa nacional de alimentación escolar del Perú, que brinda servicio alimentario con complemento educativo a niños matriculados en instituciones educativas públicas del nivel inicial y primaria (y de secundaria, en el caso de las poblaciones de la Amazonía peruana). El objetivo de PNAEQW es garantizar el servicio alimentario a los niños durante todos los días del año escolar, de modo que mejoren su atención en las clases, incrementen su asistencia y permanencia en la escuela y mejoren sus hábitos alimenticios. El PNAEQW es un programa universal, es decir, busca atender a todos los niños de todas las instituciones públicas de inicial y primaria a nivel nacional. Actualmente atiende a más de 3.8 millones de estudiantes.

El PNAEQW funciona a partir de un modelo de cogestión del servicio alimentario que comprende las siguientes fases: (1) la planificación del menú escolar; (2) el proceso de compra; y (3) la gestión del servicio alimentario. La planificación de menú escolar se realiza a partir de las recomendaciones nutricionales del Centro Nacional de Alimentación y Nutrición (CENAN). Por su parte, el proceso de compra está a cargo del Comité de Compras (CC) y su función es adquirir las raciones y productos para la prestación del servicio alimentario. Finalmente, la gestión del servicio alimentario está a cargo del Comité de Alimentación Escolar (CAE) y su proceso comprende las siguientes etapas: (1) recepción de productos; (2) almacenamiento de productos; (3) preparación de alimentos; (4) servido de alimentos; (5) distribución de alimentos y (6) consumo de alimentos.

Este modelo cogestionado funciona con la participación articulada y cooperación de tres actores principales: la comunidad local, el sector público y el sector privado. Estos actores se organizan tanto en la conformación del CC como del CAE. El primero está conformado por un subprefecto, un representante de la municipalidad, un representante de la Red de Salud y dos representantes de los padres de familia; mientras que el CAE está compuesto por un director (presidente), un docente (secretario) y tres representantes de los padres de familia (vocales).

Actualmente, el PNAEQW entrega el servicio alimentario a partir de dos modalidades: productos y raciones. Los productos son preparados en las comunidades para la provisión de desayunos y almuerzos y se componen de un bebible (a base de granos, como la kiwicha o la quinua) y un sólido (galletas o cereales acompañados de un alimento de origen animal como pollo, carne, etc.). Por otro lado, las raciones son alimentos listos para consumir que solo se entregan como desayuno. Al igual que los productos, consisten en un bebible, en este caso procesado, y un sólido. El bebible puede ser leche enriquecida o un cereal con leche y el sólido usualmente consiste en pan o galletas hechas de cereales locales o *snacks* saludables.

Si bien el PNAEQW es un programa universal, la entrega del servicio se realiza de manera diferenciada. A los alumnos de las escuelas de los distritos más pobres (quintiles 1 y 2) se les entrega dos raciones de alimentos (desayuno y almuerzo), mientras que a los alumnos de las escuelas de los distritos más ricos (quintiles 3, 4 y 5) solo se les entrega una ración (desayuno). Una particularidad es que el servicio es homogéneo a nivel distrital: todas las escuelas dentro

de un mismo distrito reciben la misma modalidad de atención (desayuno o desayuno + almuerzo). Por otro lado, si bien el servicio se entrega de manera diferenciada según el nivel de pobreza del distrito al que pertenece cada escuela (dimensión socioeconómica), no toma en cuenta la severidad del principal problema que el PNAEQW intenta resolver: la inadecuada ingesta de alimentos.

En este contexto, surge la necesidad de priorizar el servicio alimentario de PNAEQW de modo que permita de cerrar las brechas de ingesta de alimentos de manera más eficiente. El presente estudio tiene como objetivo la construcción de un índice que permita dicha priorización.

2. CONTEXTO DEL ESTUDIO

Como se mencionó en la sección anterior, este estudio surge en un contexto en el que se ha visto necesaria la priorización del servicio alimentario que ofrece el PNAEQW, o por lo menos, de la diferenciación de su servicio tomando en consideración la dimensión nutricional. Esta necesidad tiene como origen la solicitud del Ministerio de Economía y Finanzas (MEF) en 2016 de revisar los criterios de priorización del PNAEQW de modo que no solo respondan a condiciones socioeconómicas, sino que incorporen aspectos nutricionales y, en ese sentido, responda al problema que intenta resolver el programa que es el déficit calórico². De acuerdo con lo señalado en la Nota Metodológica, lo anterior se refuerza con los resultados que las distintas evaluaciones que se han realizado en torno a la implementación del programa, donde se destaca la necesidad de ajustar o rediseñar algunas dimensiones de este, entre ellos, la oportunidad y la relevancia del servicio entregado.

Por ejemplo, una evaluación de los procesos de entrega de los servicios que ofrece el PNAEQW realizada por el MIDIS en 2016 encontró que los niños que pertenecen a los quintiles más ricos toman desayuno en sus hogares; y que sus padres los envían con una lonchera adicional³. Por tanto, estos niños estarían consumiendo entre dos y tres comidas antes del almuerzo (incluyendo lo brindado por el PNAEQW). El estudio realizado por la Oficina del Programa Mundial de Alimentos en el Perú (en adelante, **PMA**)⁴ en 2016 para el PNAEQW también encontró que el 78% de los estudiantes tomaban desayuno en el hogar. En esa misma línea, otro estudio realizado por Macroconsult e Instituto Cuánto (2016) encontró que el 84% del total de beneficiarios del PNAEQW tomaba desayuno en su hogar, por lo que los alimentos brindados por el PNAEQW eran consumidos a modo de lonchera y no como una comida principal⁵. No obstante, un estudio realizado por AC Pública (2016), señala que, en los quintiles más pobres, el desayuno que consumen los niños en su hogar consiste solo en mate y caldo, por lo que llegan a la escuela dispuesto a consumir el desayuno proporcionado por el programa⁶.

² Ver Informe de Verificación del Programa Presupuestal “Programa Nacional de Alimentación Escolar” elaborado por la Dirección General de Presupuesto Público del MEF, donde se detallan las principales observaciones y recomendaciones que deben incorporarse en el Anexo N°2 del PP 0115 correspondiente al ejercicio 2017; citado en la Nota Metodológica.

³ Ver reporte elaborado por la Dirección General de Seguimiento y Evaluación del MIDIS en 2016 denominado “Evaluación del proceso de provisión del servicio alimentario en instituciones educativas del Programa Nacional de Alimentación Escolar Qali Warma”; citado en la Nota Metodológica.

⁴ Ver informe encargado por el PMA en 2016 denominado “Caracterización del consumo alimentario de los usuarios del Programa Nacional de Alimentación Escolar Qali Warma y la contribución de la ración entregada por el programa a escolares en Lima Metropolitana 2016”; citado en la Nota Metodológica.

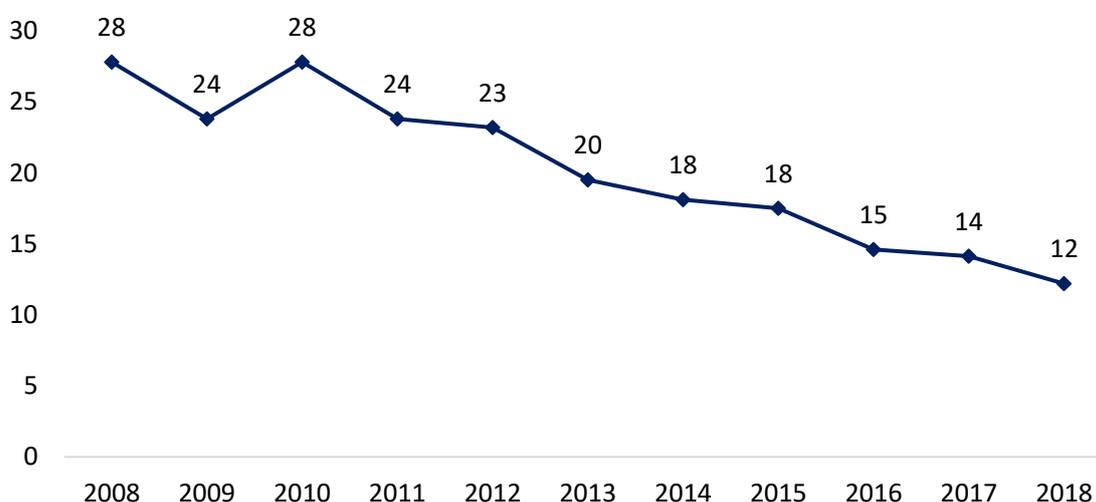
⁵ Ver Informe Final del estudio elaborado por Macroconsult e Instituto Cuánto en 2016 denominado “Elaboración del estudio de percepción de la calidad de los servicios de los programas sociales del MIDIS”; citado en la Nota Metodológica.

⁶ Ver Informe Final del estudio elaborado por AC Pública en 2016 denominado “Evaluación del proceso de provisión del servicio alimentario en instituciones públicas del Programa Nacional de Alimentación Escolar Qali Warma”; citado en la Nota Metodológica.

Aunado a lo anterior se tiene que el Ministerio de Salud, como parte del Plan de Salud Escolar (2015), encontró que el 14.1% de los estudiantes de instituciones públicas de nivel inicial y primario a nivel nacional presentan problemas de exceso de peso⁷. A ello, se suman los resultados encontrados por la misión del PMA en 2014 sobre la composición nutricional de las raciones provistas por el PNAEQW⁸. Si bien la misión manifestó que, en términos generales, los desayunos y almuerzos estaban balanceados, también observó que el porcentaje de energía proveniente del azúcar en los desayunos era mayor del recomendado para personas sin sobrepeso y que los almuerzos eran bajos en micronutrientes.

La evidencia presentada hasta el momento sugiere revisar las modalidades de atención que ofrece el programa, especialmente en un contexto de doble carga nutricional como el peruano. Este término hace alusión a un contexto de “transición nutricional” en la que la reducción de los problemas de desnutrición infantil ocurre a la par que el aumento de los problemas asociados al sobrepeso (Black et al, 2013). Si bien en el Perú la información disponible para caracterizar el estado nutricional de los niños mayores de cinco años es limitada, se cuentan con algunos hechos estilizados. Por un lado, sabemos que, en los últimos diez años, la desnutrición crónica infantil (en niños menores de 5 años) se ha reducido dramáticamente, pasando de 28% en 2008 a menos de la mitad en 2018 (12%) (ver **Gráfico 1**). Por otro, de acuerdo a la información del Instituto Nacional de Salud (INS), el sobrepeso y la obesidad alcanzaría a 3 de cada 10 niños de 5 a 9 años (ver **Gráfico 2**), y en la misma cantidad a los niños y adolescentes de 10 a 19 años (ver **Gráfico 3**).

Gráfico 1. Prevalencia de la desnutrición infantil crónica, 2008-2018
(Porcentaje)

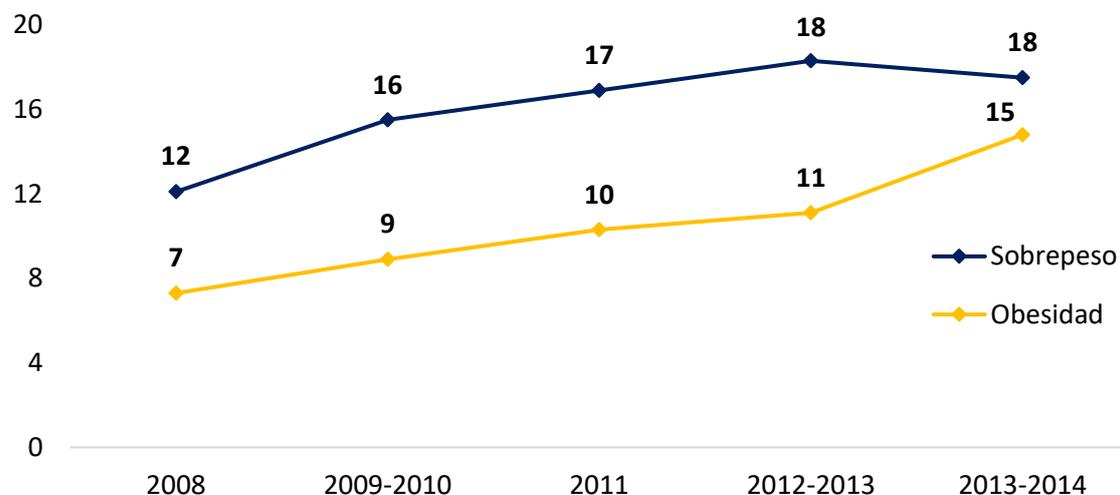


Fuente: Encuesta Demográfica y de Salud Familiar (ENDES) 2008 – 2018. **Elaboración propia.**

⁷ Ver los resultados de la evaluación del Plan de Salud Escolar en 2015 elaborado por el Ministerio de Salud; citado en la Nota Metodológica.

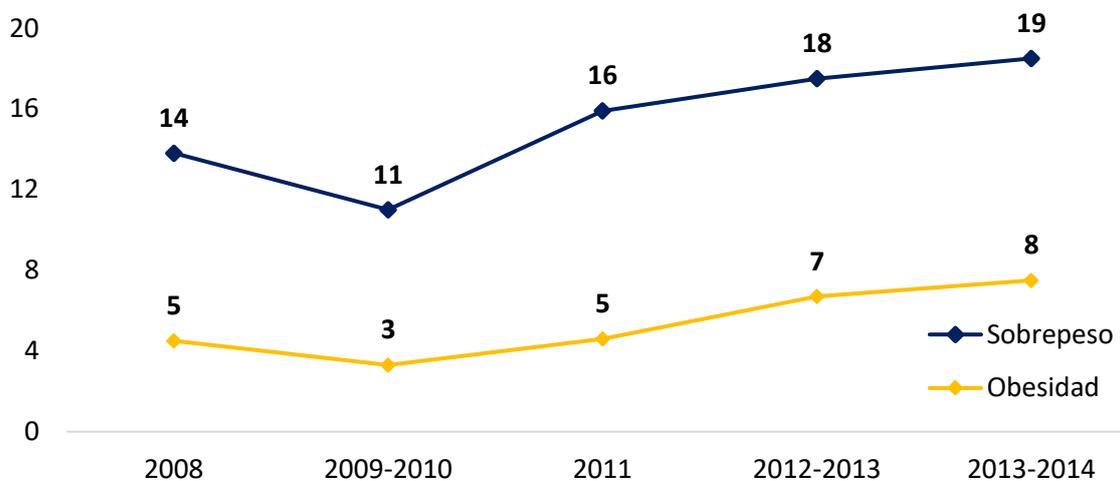
⁸ Ver el reporte elaborado por el Programa Mundial de Alimentos en octubre de 2014 denominado “*Back to office report – Peru School Feeding Review Mission*”; citado en la Nota Metodológica.

Gráfico 2. Evolución de la prevalencia del sobrepeso y obesidad infantil en niños de 5 a 9 años, 2007-2014
(Porcentaje)



Fuente: Observatorio de Nutrición y Estudio del Sobrepeso y Obesidad – Instituto Nacional de Salud (INS).
Elaboración propia.

Gráfico 3. Evolución de la prevalencia del sobrepeso y obesidad infantil en niños de 10 a 19 años, 2007-2014



Fuente: Observatorio de Nutrición y Estudio del Sobrepeso y Obesidad – Instituto Nacional de Salud (INS).
Elaboración propia.

Por ese motivo, se justifica la necesidad del diseño de una metodología costo efectiva que brinde un criterio de priorización para el servicio alimentario provisto por el PNAEQW que permita cumplir con su objetivo: mejorar la inadecuada ingesta de alimentos y la promoción de hábitos de alimentación saludables. El presente estudio tiene como objetivo la construcción de un índice que permita dicha priorización.

3. OBJETIVOS DEL ESTUDIO

El objetivo del presente estudio es elaborar y validar el índice de priorización para la prestación del servicio alimentario del PNAEQW en el nivel de primaria. Para ello, se definieron los siguientes objetivos específicos:

- Armonizar las bases de datos necesarias para la construcción del índice de priorización.
- Validar el índice de priorización construido a partir de las bases de datos disponibles con un índice construido a partir de información primaria.
- Elaborar el mapa de priorización para la prestación del servicio alimentario de PNAEQW.
- Desarrollar el protocolo para la actualización anual del índice de priorización de PNAEQW.

4. RESUMEN METODOLÓGICO

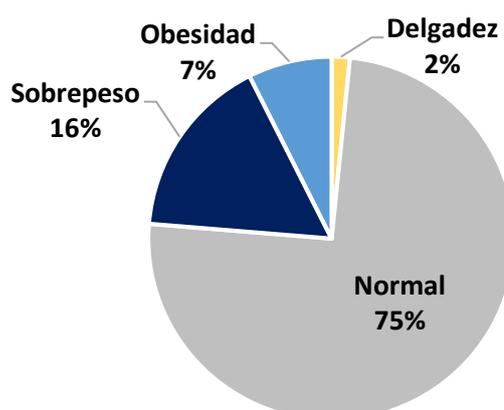
En esta sección se presenta un resumen de la metodología empleada para la construcción de los índices de priorización; el detalle de esta se puede encontrar en el **Anexo 12.1** (Anexo Metodológico). Las actividades desarrolladas en el marco de este estudio pueden ser organizadas en tres grandes actividades: (i) la armonización de las bases de datos necesarias para la construcción de los índices; (ii) la construcción de los índices de priorización; y (iii) la clasificación de las escuelas en las categorías definidas.

La primera actividad, la armonización de las bases de datos, estuvo a cargo del INEI. Se armonizaron cuatro bases de datos: el padrón de escuelas atendidas por el PNAEQW en el año 2019 (en adelante, **padrón PNAEQW**), el padrón de alumnos del Sistema de Información de Apoyo a la Gestión de la Institución Educativa (SIAGIE) 2017 (en adelante, **SIAGIE 2017**), el padrón de población y vivienda de los Censos Nacionales del 2017 (en adelante, **Censo 2017**), y el Plan de Salud Escolar (en adelante, **PSE**). Las actividades del equipo consultor durante esta fase se limitaron a revisar la consistencia de la información recibida y a realizar la limpieza de las bases de datos. El universo de información viene dado por el padrón del PNAEQW, donde se registran 29 mil 486 escuelas públicas de nivel primario atendidas en la actualidad que benefician a cerca de 2.6 millones de alumnos. Como resultado del empate realizado, se cuenta con información socioeconómica del Censo 2017 del 98% de escuelas (28,742 escuelas) y del 72% de alumnos (1,804,232 alumnos); y con información de talla y peso del 91% de escuelas (26,787 escuelas) y del 42% de los alumnos (1,038,754).

La segunda actividad, referida a la construcción de los índices de priorización, se realizó de manera diferenciada para cada dimensión del índice (socioeconómica y nutricional), aun cuando la metodología fue la misma en ambos casos. En ambos casos, el procedimiento fue el siguiente: primero, se construyeron los indicadores a nivel de alumnos sobre la base de la información existente y luego se agregaron a nivel de escuela. En el caso del índice socioeconómico, se construyeron dos índices, uno a partir de la tasa de pobreza según el método de las Necesidades Básicas Insatisfechas (en adelante, **NBI**) de los alumnos de la escuela, y otro a partir de un índice de riqueza construido utilizando el puntaje promedio del primer componente principal de un ejercicio de Análisis de Componentes Principales (en adelante, **PCA**). En el caso del índice nutricional, se utilizó una de las sugerencias de la Nota Metodológica, el Índice de Masa Corporal (en adelante, **IMC**), a partir del cual se calculó la carga de malnutrición de cada escuela⁹. Como se observa en el **Gráfico 4**, esta última se concentra fundamentalmente en el sobrepeso y la obesidad por lo que ese fue el indicador seleccionado junto con el equipo técnico del PMA. De acuerdo con los resultados del PSE, la delgadez y la delgadez extrema –la otra cara de la malnutrición–, afectan al 2% de los estudiantes evaluados, por lo que esta condición no constituye un indicador relevante para efectos de la priorización.

⁹ La sugerencia de la nota metodológica es una combinación entre el IMC y la tasa de desnutrición crónica. Junto con el PMA se decidió optar solo por el IMC debido al grupo etario considerado en el estudio y la disponibilidad de información en el PSE usado en el análisis.

Gráfico 4. Estado nutricional de los alumnos de primaria según el PSE, 2015-2016

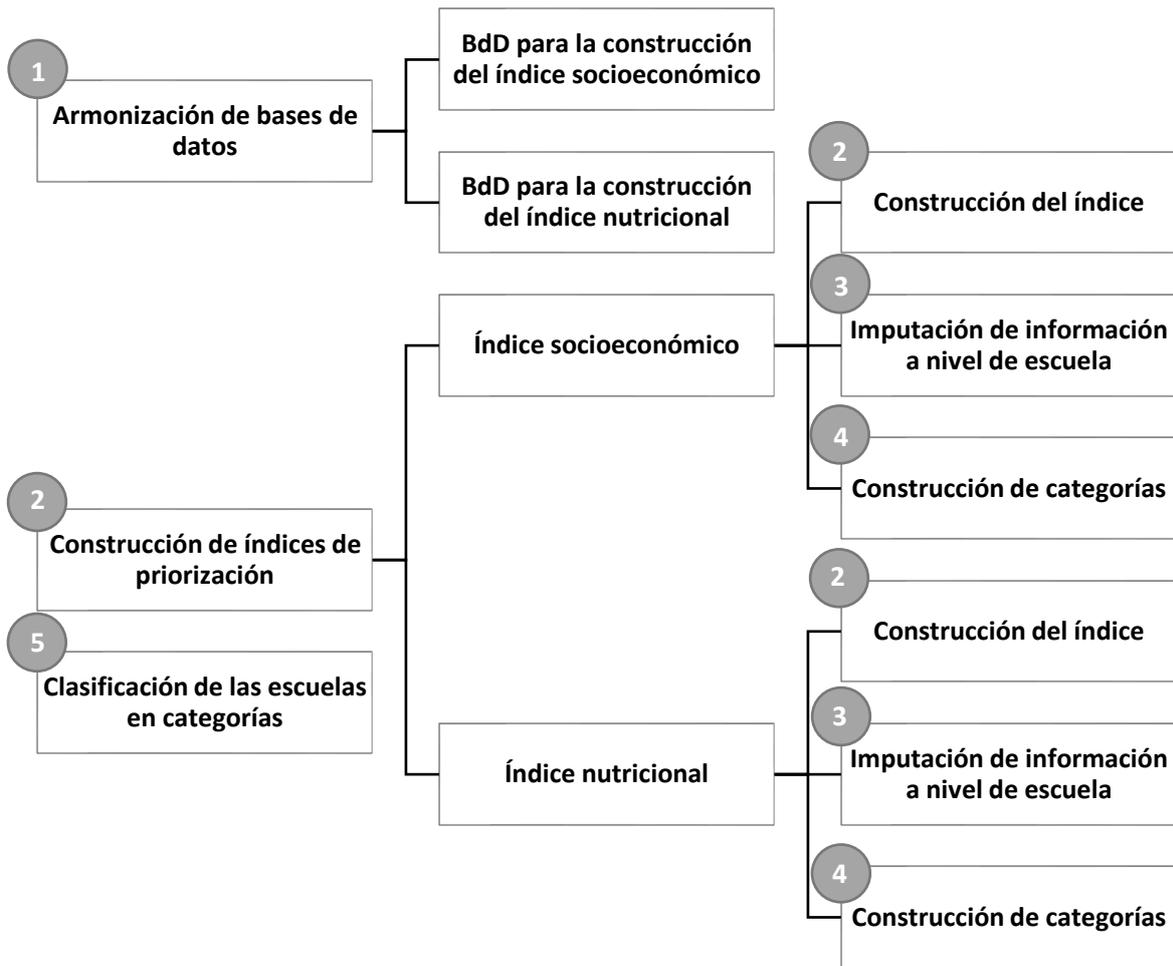


Fuente: Base de datos armonizada. **Elaboración propia.**

A continuación, se estimó la proporción de escuelas que requiere algún procedimiento de imputación. En el caso del índice socioeconómico, se requirió imputar información al 6% de escuelas que atiende el PNAEQW (1,934 escuelas) a partir del valor promedio de las escuelas vecinas más cercanas (según distancia geográfica); mientras que en el caso del índice nutricional fue necesario imputar al 32% de las escuelas a partir de una estrategia que, en una primera etapa, utiliza técnicas de predicción de *machine learning* y, en una segunda etapa, utiliza la misma estrategia de cercanía de escuelas vecinas que en el índice socioeconómico. Con la data ya imputada, se realizó el agrupamiento de las escuelas bajo cada índice. De manera *a priori*, se establecieron tres grupos de interés para el índice socioeconómico (necesidad alta, media y baja) y tres para el índice nutricional (prevalencia de sobrepeso y obesidad alta, media y baja). En el estudio se presentan tres ejercicios de agrupamiento para el caso del índice socioeconómico (dos aproximaciones utilizando quintiles de riqueza y un análisis por conglomerados) y uno para el nutricional (a partir de umbrales definidos por el equipo técnico del PMA considerando los umbrales establecidos por la Organización Mundial de la Salud (OMS) para identificar a la anemia como un problema de salud pública, así como la distribución de la carga de malnutrición de las escuelas).

Finalmente, en la tercera actividad, se integraron ambas clasificaciones dando origen al resultado final del estudio. En el **Gráfico 5** se presenta el resumen de la metodología implementada.

Gráfico 5. Metodología para la priorización del PNAEQW



Elaboración propia.

5. RESULTADOS DE LA PRIORIZACIÓN

5.1. RESULTADOS DE LA INTEGRACIÓN: ÍNDICE DE PRIORIZACIÓN

El objetivo del presente estudio es clasificar a las escuelas de acuerdo a la condición socioeconómica de los alumnos y su estado nutricional sobre la base de índices de priorización (uno socioeconómico y otro nutricional). A partir del índice socioeconómico, se clasificó a las escuelas en tres grupos: alta, media y baja necesidad; mientras que, a partir del índice nutricional, se clasificaron a las escuelas en los siguientes grupos: prevalencia de sobrepeso/obesidad alta, media y baja. El cruce entre ambas clasificaciones da origen a nueve posibles categorías que se corresponden con distintas modalidades de atención. Estas últimas fueron sugeridas por el equipo técnico del PMA sobre la base de la experiencia internacional y la asistencia técnica que el Programa Mundial de Alimentos brinda a los programas de alimentación escolar en el mundo; y son las que se presentan en la siguiente tabla. Cabe precisar que estas modalidades de atención no solo tienen implicancias en la diferenciación del tipo de menú que ofrecerá el programa considerando la carga socioeconómica y nutricional de las escuelas, sino también sobre las características del menú entregado (tipos de combinaciones, grupos de alimentos y proporciones) así como del paquete educativo y comunicacional que se entrega a las escuelas.

Tabla 4. Categorías sugeridas

Categorías del índice nutricional	Categorías del índice socioeconómico		
	Necesidad alta	Necesidad media	Necesidad baja
Sobrepeso/obesidad alta	Desayuno	Refrigerio	Refrigerio
Sobrepeso/obesidad media	Desayuno + almuerzo	Desayuno	Refrigerio
Sobrepeso/obesidad baja	Desayuno + almuerzo	Desayuno	Refrigerio

Elaboración propia.

Las categorías anteriores se pueden resumir en tres modalidades de atención. La modalidad 1 corresponde a desayunos, y se entregaría a los alumnos de las escuelas clasificadas como de alta necesidad y que tienen alto sobrepeso/obesidad, y a las escuelas clasificadas como de media necesidad y sobrepeso/obesidad media y baja. La modalidad 2 corresponde a desayunos y almuerzos, y se entregaría a los alumnos de las escuelas clasificadas como de alta necesidad y sobrepeso/obesidad media y baja. Finalmente, la modalidad 3 corresponde a refrigerios, y se entregarían al resto de escuelas. Esta modalidad constituiría la modalidad “base”.

Gráfico 6. Modalidades de atención sugeridas por el PMA



Elaboración propia.

Como se detalla en el anexo metodológico de este estudio (**ver Anexo 12.1**), en el caso del índice socioeconómico, se desarrollaron dos metodologías para la construcción del índice y tres formas de realizar la clasificación de escuelas; mientras que, en el caso del índice nutricional, se desarrolló una sola metodología tanto para la construcción del índice como para el agrupamiento de escuelas. En breve, se han desarrollado seis formas de clasificar a las escuelas. Entre las metodologías para la construcción del índice socioeconómico, el equipo consultor sugiere utilizar el índice de riqueza calculado a partir del análisis de componentes principales toda vez que este indicador se encuentra altamente correlacionado con el gasto per cápita. Luego, entre las metodologías de clasificación, recomendamos utilizar el agrupamiento a través del análisis de conglomerados (k-medias) en la medida que reduce la discrecionalidad en la construcción de grupos.

Si bien en los archivos digitales de este estudio, se presentan las seis alternativas para clasificar a las escuelas (dos metodologías de construcción del índice: NBI y PCA, cada una con tres posibles agrupamientos), a continuación, se presentan los resultados de aquella que sugerimos utilizar, la cual se construye sobre la base del índice de riqueza (PCA) que agrupa a las escuelas siguiendo la metodología del análisis por conglomerados. En la **Tabla 5** se presenta la cantidad de escuelas y de alumnos en cada categoría y la expresión de ambas cantidades como proporción de la cantidad total de escuelas y alumnos, respectivamente. Como se observa, la categoría más frecuente es la de escuelas con alta necesidad y medio sobrepeso/obesidad (10,192 escuelas), seguido por la de alta necesidad y bajo sobrepeso/obesidad (6,887 escuelas). No obstante, al tratarse de escuelas pequeñas, no se corresponden con un gran número de alumnos. La mayor parte de alumnos se encuentra concentrado en la categoría baja necesidad y alto sobrepeso/obesidad (1.2 millones de alumnos), a pesar de que, en términos de escuelas, no representan una cantidad importante.

Tabla 5. Distribución de escuelas y de alumnos según índices de priorización*

Clasificación según índice nutricional	Clasificación según índice socioeconómico			Total
	Necesidad alta	Necesidad media	Necesidad baja	
Cantidad y porcentaje de escuelas				
Sobrepeso/obesidad alta	1,641 6%	2,232 8%	3,101 11%	6,974 24%
Sobrepeso/obesidad media	10,192 35%	3,816 13%	258 1%	14,266 48%
Sobrepeso/obesidad baja	6,887 23%	1,237 4%	122 0%	8,246 28%
Total	18,720 63%	7,285 25%	3,481 12%	29,486 100%
Cantidad y porcentaje de alumnos				
Sobrepeso/obesidad alta	67,699 3%	319,742 12%	1,160,855 45%	1,548,296 60%
Sobrepeso/obesidad media	383,265 15%	363,162 14%	81,811 3%	828,238 32%
Sobrepeso/obesidad baja	172,723 7%	33,590 1%	7,223 0%	213,536 8%
Total	623,687 24%	716,494 28%	1,249,889 48%	2,590,070 100%

*Nota: Los porcentajes están expresados respecto al total de escuelas y de alumnos atendidos por el PNAEQW.

Fuente: Base de datos armonizada. **Elaboración propia.**

Dado que la propuesta de modalidades de atención es la misma para las escuelas sobrepeso/obesidad media y baja, el contenido de la **Tabla 5** puede resumirse en seis categorías tal como se presenta en la **Tabla 6**.

Tabla 6. Distribución de escuelas y de alumnos según índices de priorización y categorías sugeridas*

Clasificación según índice nutricional	Clasificación según índice socioeconómico			Total
	Necesidad alta	Necesidad media	Necesidad baja	
Cantidad y porcentaje de escuelas				
Sobrepeso/obesidad alta	1,641 6%	2,232 8%	3,101 11%	6,974 24%
Sobrepeso/obesidad media y baja	17,079 58%	5,053 17%	380 1%	22,512 78%
Total	18,720 63%	7,285 25%	3,481 12%	29,486 100%
Cantidad y porcentaje de alumnos				
Sobrepeso/obesidad alta	67,699 3%	319,742 12%	1,160,855 45%	1,548,296 60%
Sobrepeso/obesidad media y baja	555,988 22%	396,752 15%	89,034 3%	1,041,774 40%
Total	623,687 24%	716,494 28%	1,249,889 48%	2,590,070 100%

*Nota: Los porcentajes están expresados respecto al total de escuelas y de alumnos atendidos por el PNAEQW.

Fuente: Base de datos armonizada. Elaboración propia.

5.2. ESTADÍSTICAS DESCRIPTIVAS: COMPARACIÓN CON LA SITUACIÓN ACTUAL

Para poner los resultados del ejercicio realizado en contexto de la distribución que actualmente maneja el programa, resulta útil observar los totales de las categorías socioeconómicas y compararlas con los equivalentes de la **Tabla 7**. Para esto, y con ánimo de hacer comparables las cifras, se ha considerado a las escuelas ubicadas en el primer quintil de pobreza del padrón del PNAEQW como aquellas de alta necesidad, a las de los quintiles 2 y 3 como de necesidad media, y a las de quintiles 4 y 5 como de necesidad alta. No se utiliza la clasificación que utiliza de modo formal PNAEQW para efectos operativos (quintiles 1 y 2 por un lado y 3, 4 y 5 por otro), porque no permite una comparación equivalente con las categorías que se propone.

Tabla 7. Clasificación actual del PNAEQW para la atención de escuelas*

Clasificación actual	Clasificación socioeconómica			Total
	Necesidad alta	Necesidad media	Necesidad baja	
Escuelas	7,306	14,306	7,874	24,486
	30%	58%	32%	100%
Alumnos	308,523	802,812	1,477,735	2,590,070
	12%	31%	57%	100%

***Nota:** Para hacer comparables las cifras, se ha considerado a las escuelas ubicadas en el quintil más pobre del padrón de PNAEQW como aquellas de alta necesidad, a las de los quintiles intermedios como de necesidad media, y los dos quintiles más ricos como de necesidad baja.

Fuente: Base de datos armonizada. **Elaboración propia.**

Por ejemplo, con la propuesta del presente estudio, se tienen 18,720 escuelas en la categoría de alta necesidad socioeconómica, que es casi 2.5 veces el número de escuelas que actualmente el programa considera en esa categoría (7,306). De manera similar, bajo el ejercicio realizado se contabilizan 623,687 usuarios de alta necesidad socioeconómica, que es poco más del doble de los que actualmente contabiliza el PNAEQW (309,523). Más contrastes hay a nivel de escuelas en la categoría de necesidad media, sin embargo. Bajo el ejercicio realizado en el estudio, se clasifican 7,285 escuelas, pero el programa contabiliza 14,306 actualmente. Solo a nivel de alumnos, las discrepancias son menores, pues contabilizamos 716,494 alumnos de necesidad media y el programa 802,812. Las diferencias emergen nuevamente en la categoría de baja necesidad, principalmente a nivel de escuela, ya que nuestra metodología clasifica a 3,481 en esta categoría y el programa contabiliza 7,874. A nivel de estudiantes, las cifras son de 1,249,889 contra 1,477,735 que tiene el programa actualmente.

Estas diferencias en la dimensión socioeconómica se explican por el nivel de agregación para la realización de la clasificación socioeconómica. En particular, PNAEQW actualmente utiliza una clasificación distrital, lo que hace que se incluyan en distritos de alta pobreza a escuelas grandes que podrían albergar a alumnos con relativamente baja necesidad y viceversa. Nuestra metodología, a diferencia, al descansar en variables agregadas a nivel de escuela, permite una clasificación más precisa.

Una segunda arista de comparación es la derivada de las modalidades de atención que surgen de la propuesta versus las que actualmente tiene el programa. Esto se presenta en la **Tabla 8**. Actualmente, PNAEQW entrega 1.9 millones de raciones de la modalidad de desayunos y 716 mil raciones de la modalidad desayuno + almuerzo. Bajo la propuesta alterativa, se sugiere que PNAEQW entregue 1.6 millones de raciones de refrigerios, 464 mil raciones de la modalidad desayuno y 556 mil raciones de la modalidad desayuno + almuerzos. Siendo así, las diferencias no implican cambios operativos muy disruptivos, lo cual sugiere que tampoco tendrán consecuencias presupuestales mayores, pues se reduce la cantidad de raciones de la modalidad compleja y costosa (desayuno + almuerzo) y se aumenta las raciones de las modalidades sencillas y baratas (refrigerio y desayuno).

Añadir una nueva modalidad de atención en la forma de refrigerios tampoco es ajeno al programa, pues, según indicación del PNAEQW, en la actualidad se pilotea este tipo de

modalidad en un número pequeño de escuelas de Villa María del Triunfo y Lima. Consideramos que este es un antecedente válido de la factibilidad de la propuesta.

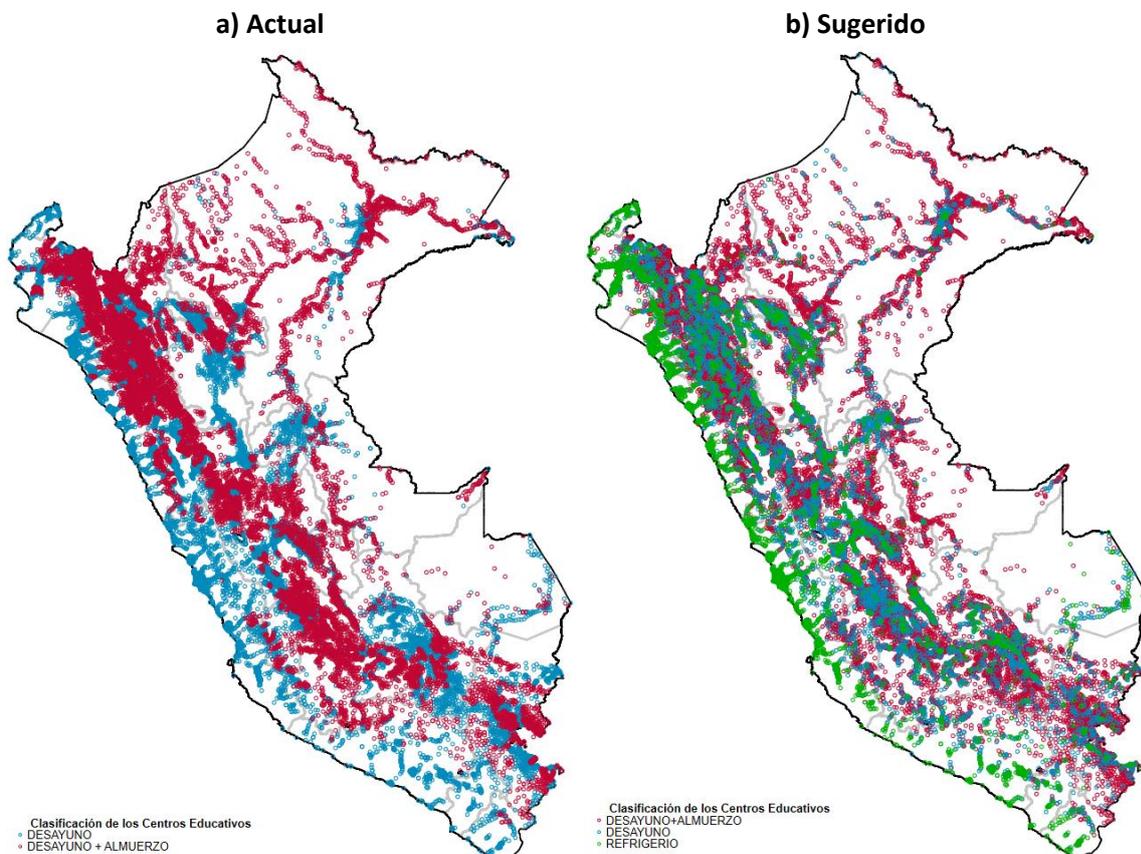
Tabla 8. Comparación de las modalidades de atención

	Modalidad 1: Desayuno	Modalidad 2: Desayuno + almuerzo	Modalidad 3: Refrigerio	Total
Propuesta				
Escuelas	6,694	17,079	5,713	29,486
Alumnos	464,451	555,988	1,569,631	2,590,070
Situación actual				
Escuelas	13,642	15,844	-	29,486
Alumnos	1,873,969	716,101	-	2,590,070

Fuente: Base de datos armonizada. **Elaboración propia.**

Finalmente, las diferencias en las modalidades de atención se presentan en el siguiente mapa. Este dice que las principales variaciones ocurren principalmente en la sustitución de desayunos por refrigerios, principalmente en la costa del país.

Gráfico 7. Distribución de las modalidades de atención a nivel escuelas

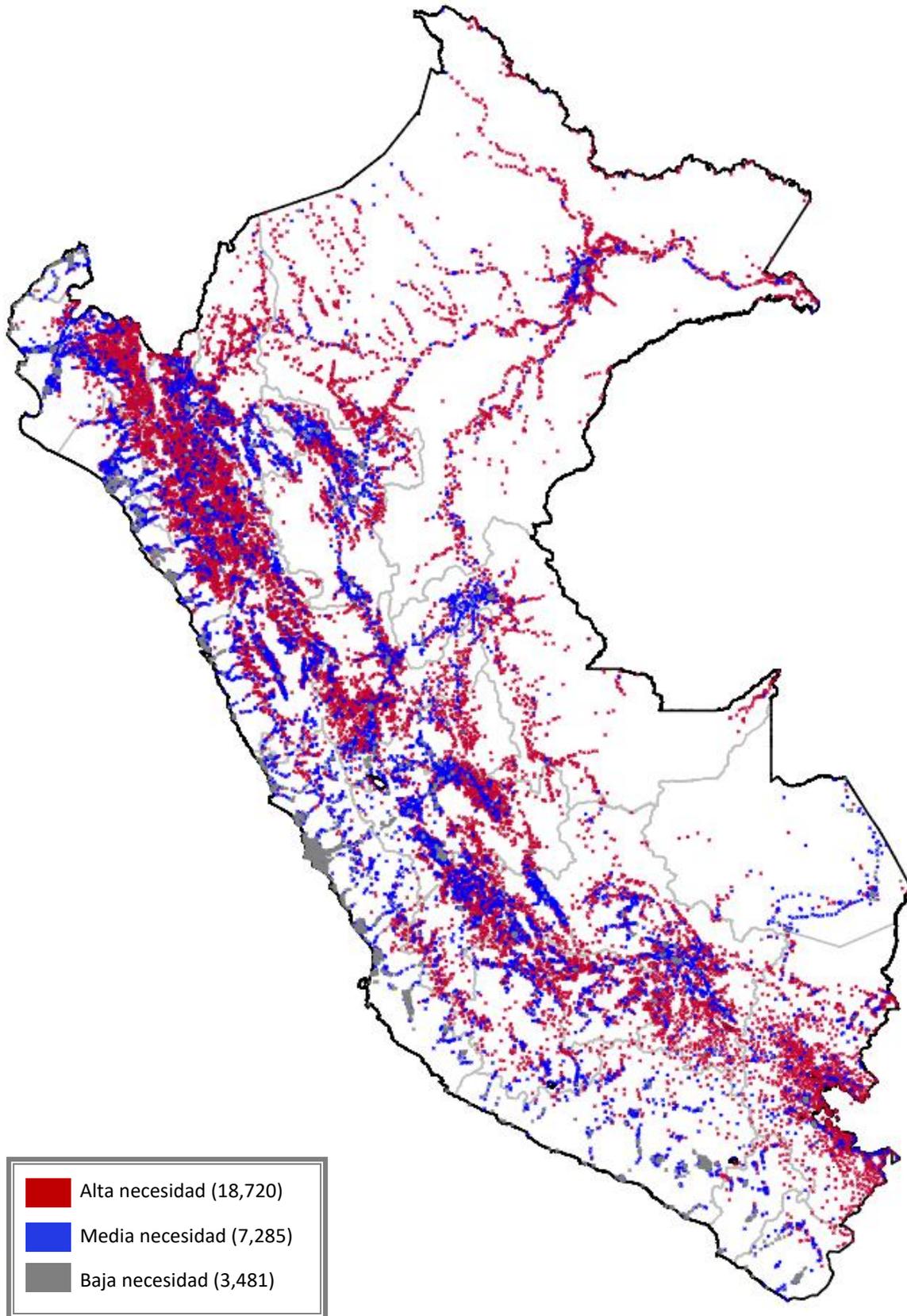


Fuente: Base de datos armonizada. **Elaboración propia.**

6. MAPAS DE PRIORIZACIÓN 2018

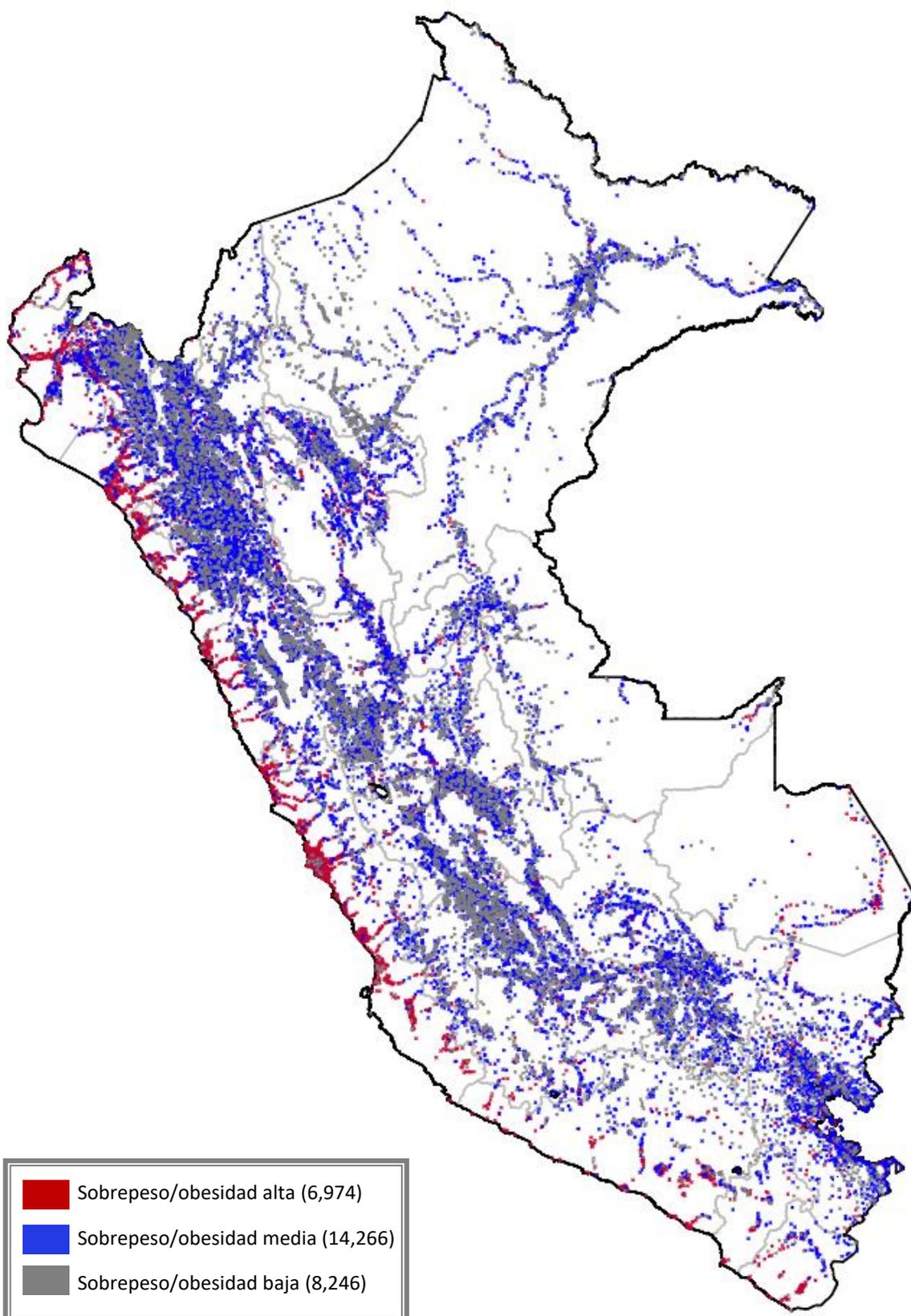
A continuación, se presentan los mapas de priorización para el año 2018. El **Mapa 1** corresponde a la clasificación de las escuelas en alta, media y baja necesidad según el índice socioeconómico seleccionado. Por su parte, en el **Mapa 2** se presenta la distribución de escuelas según las clasificaciones del índice nutricional (alto, medio y bajo sobrepeso/obesidad). Finalmente, en el **Mapa 3** se muestra la distribución de las escuelas según las nuevas modalidades de atención que se derivan del cruce de ambos índices. Cabe recordar que la modalidad 1 corresponde a la entrega de desayuno, la modalidad 2 a desayuno y almuerzo, y la modalidad 3 a refrigerios.

Mapa 1. Distribución de escuelas según índice socioeconómico



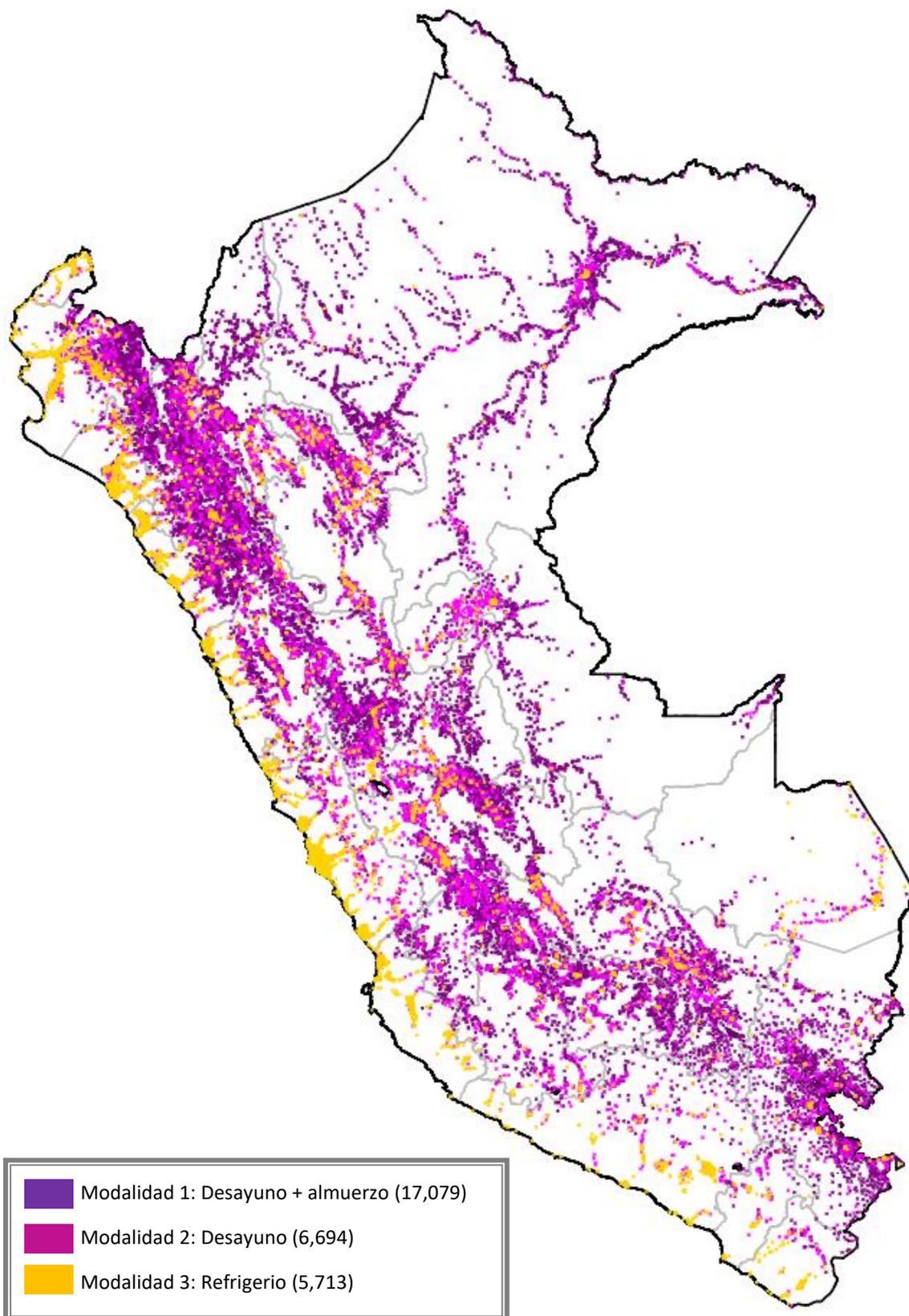
Fuente: Base de datos armonizada. Elaboración propia.

Mapa 2. Distribución de escuelas según índice nutricional



Fuente: Base de datos armonizada. Elaboración propia.

Mapa 3. Mapa de priorización del PNAEQW 2018



Fuente: Base de datos armonizada. Elaboración propia.

7. LIMITACIONES DEL ESTUDIO

Se han identificado cuatro limitaciones de este estudio:

- 1. La información utilizada para la construcción de los índices de priorización es la mejor aproximación que encontramos, pero sigue siendo una aproximación.** En el caso del índice socioeconómico, se utilizó un índice de riqueza construido a partir de un análisis de componentes principales que, si bien se encuentra altamente correlacionado con el gasto y, en ese sentido, constituye una buena aproximación a la capacidad de gasto de los hogares de los alumnos, no es la misma variable. En el caso del índice nutricional, se utilizó la prevalencia de sobrepeso/obesidad en las escuelas medido a partir de la talla y peso de los alumnos. No obstante, la información utilizada corresponde al PSE de los años 2015 – 2016 (aproximadamente de tres años atrás). Si bien en el estudio hemos validado que la data de hace tres años se encuentra altamente correlacionada con la información actual, continúa siendo una aproximación al estado nutricional actual de los niños.
- 2. La información utilizada para la construcción de los índices de priorización no fue información completa.** Para la construcción del índice socioeconómico se utilizó data censal. No obstante, luego del emparejamiento de las bases de datos, solo se logró recuperar información del 72% de alumnos, por lo cual se tuvo que imputar información al 6% de escuelas. Por su parte, para la construcción del índice nutricional se utilizó la información del PSE de los años 2015-2016. Si bien la intención del PSE era recoger información censal, en la práctica, solo recogió información del 70% de los estudiantes atendidos por PNAEQW. Esto, aunado a la pérdida de información producto del empate de las bases de datos, ocasionó que solo se logre recuperar la información de la talla y el peso del 42% de los alumnos y se requiera imputar información al 32% de las escuelas.

La pérdida de información relacionada a la ausencia de un sistema integrado de información en el sector público no fue un tema menor en este estudio. Por un lado, la solicitud y la armonización de las bases de datos fue un proceso largo y engorroso, que tomó cerca de medio año en completarse. Por otro lado, el emparejamiento de cada base de datos a partir del DNI y de los datos personales de los alumnos tuvo como consecuencia pérdidas de información que no hubieran ocurrido si la información de las instituciones involucrada se hubiera encontrado integrada

- 3. Si bien la metodología implementada para la reconstrucción de datos faltantes fue la mejor que encontramos para superar las brechas de información, tiene un margen de error.** Como ya se mencionó, los requerimientos de imputación de este estudio fueron bastante elevados, especialmente en el caso del índice nutricional. Si bien el procedimiento realizado para la imputación permite superar exitosamente la brecha de información faltante y el resultado de nuestras estimaciones es consistente, cualquier metodología de construcción de datos siempre puede tener errores de precisión, más aún cuando el objetivo es generar información puntual a nivel de escuelas.

8. COMPARACIÓN CON OTROS ESTUDIOS Y PROGRAMAS SIMILARES

El estudio realizado tiene una larga tradición en la literatura de estimación de focalización y asignación de programas sociales, mucha de la cual descansa en la construcción de indicadores en micro espacios (Elbers et al. 2002, 2003). En general, los programas de alimentación escolar se focalizan, priorizan y asignan en función a información desagregada que condicione modalidades de atención diferenciadas.

La revisión realizada al respecto entre algunos países de la región ha permitido observar que los programas de alimentación escolar tienen características relativamente similares, y, en general, comparten el objetivo de complementar la alimentación recibida por los niños en sus hogares además de promover la asistencia escolar de las poblaciones vulnerables. Sin embargo, los criterios de priorización difieren en el tipo de información disponible. La **Tabla 9** brinda detalles sobre este punto.

Por un lado, en el Perú Qali Warma es un programa universal que proporciona alimentos a todas las escuelas públicas primarias del país. En virtud a ello, opera con un criterio de priorización orientado a atender con las modalidades de atención de desayunos y almuerzos a todas las escuelas ubicadas en los distritos de los dos quintiles más pobres según el Mapa de Pobreza 2013 construido por el INEI, y sólo con desayunos a las escuelas de los demás distritos. Operativamente, para efectos de la priorización sólo observa una variable de asignación (la tasa de pobreza), que tiene un nivel de agregación supra-escuela (el distrito).

El presente estudio propone, manteniendo el carácter universal del programa, un sistema de priorización diferenciado a nivel de escuelas de un mismo distrito, haciendo uso de criterios de priorización basados en dos dimensiones: socioeconómicas y nutricionales. Además, se plantean tres modalidades de atención, que responden a cada grupo de escuelas definidas a partir de las dos dimensiones anteriores: desayuno, desayuno y almuerzo, y refrigerio. Operativamente, el programa significa un cambio en la unidad de priorización (de distritos a escuelas), en las modalidades de atención y, por tanto, en los contenidos calóricos de las raciones.

El Programa de Alimentación Escolar (PAE) de Chile tiene criterios de asignación similares al actual, aunque contruidos a partir de indicadores y unidades diferentes. En particular, este programa es focalizado entre los alumnos que se encuentren dentro del 60% de las familias más vulnerables o con mayor desventaja socioeconómica del país. Se asigna a nivel individual, basado en el Registro Social de Hogares¹⁰, que es un sistema de información pormenorizado que utiliza datos autoreportados por los hogares en combinación con información de otras instituciones públicas del país. Este programa ofrece diferentes modalidades de atención que toma en cuenta las características socioeconómicas particulares de los estudiantes, información que recogen a partir de las encuestas entregadas por las escuelas y el Registro Social de Hogares.

¹⁰ El Registro Social de Hogares es un registro de información construido con datos aportados por el hogar en la Ficha de Protección Social o ficha social y las bases de datos que posee el Estado entre las que se cuenta información del Servicio de Impuestos Internos (SII), Registro Civil, Administradora del Fondo de Cesantía (AFC), Instituto de Previsión Social (IPS), Superintendencia de Salud y Ministerio de Educación entre otros.

Asimismo, en ese país el PAE contempla la provisión diferenciada de contenido calórico según los grupos de edad a los que pertenezcan los alumnos (JUNAEB, 2019).

Por otro lado, el Programa de Alimentación Escolar (PAE) de Colombia es universal, pero se prioriza a partir del ámbito de residencia (urbana, rural) y del nivel de vulnerabilidad de la población (comunidades étnicas, víctimas del conflicto armado, vulnerabilidad económica y nutricional). La unidad de priorización es la escuela. En una primera etapa se priorizan las escuelas de zonas rurales o que atienden exclusivamente a poblaciones étnicas, y a las escuelas de las zonas urbanas que tengan alta concentración de población en situación de vulnerabilidad (medida como puntajes en el Sistema de Identificación de Potenciales Beneficiarios de Programas Sociales - SISBEN). Además, se prioriza a las escuelas de preescolar, primaria, y primaria para grupos étnicos o con alta proporción de población víctima del conflicto armado. En la segunda etapa, se selecciona a los beneficiarios de cada institución educativa. En el caso de las escuelas rurales y que atienden a población étnica la atención cubre al total de los alumnos. En las escuelas urbanas se atiende al total de alumnos, dando preferencia a los niños provenientes de comunidades étnicas, víctimas del conflicto armado, clasificados como vulnerables por el SISBEN, y se prioriza los grados inferiores por encima de los superiores. La asignación de modalidades de atención se realiza teniendo en cuenta las necesidades nutricionales según grupos de edad (MEN, 2013).

El Programa Nacional de Alimentación Complementaria Escolar (PNACE) de Bolivia corresponde a un programa en desarrollo en el que los mecanismos de priorización no están homogenizados debido a que la gestión es descentralizada y a nivel municipal. Este programa se creó con el fin de universalizar el acceso a los desayunos escolares y uniformizar los criterios de calidad y cantidad en la provisión de las raciones a lo largo del territorio boliviano. Esto dado que su predecesora, la Alimentación Complementaria Escolar (ACE) se desarrollaba de forma muy heterogénea entre municipios (Ministerio de Educación, 2015). Es por eso que, en el caso de Bolivia, el programa de alimentación escolar se encuentra en proceso de un desarrollo gradual. La **Tabla 9** presenta la descripción de la provisión de alimentos basada en los lineamientos del PNACE. Como se observa, este programa ofrece hasta cuatro modalidades de atención que se priorizan en función de la clasificación de los municipios. Esta clasificación la realiza el Ministerio de Desarrollo Rural y Tierras (MDRyT) hecho con base en el Mapa de Vulnerabilidad a la Inseguridad Alimentaria de 2012. Todas las escuelas públicas dentro de municipio reciben la misma modalidad de atención.

Finalmente, la experiencia del Programa de Alimentación Escolar (PAE) en Ecuador se inició en 1999. A partir del 2000 dejó de ser focalizado con el propósito de universalizarlo. A partir del 2010, el PAE es incluido dentro del Plan Nacional de Desarrollo para el Buen Vivir 2009 – 2013. Este programa tiene dos modalidades de atención cuya asignación se basa en la clasificación de los centros educativos. Aquellas instituciones residenciales reciben desayuno, almuerzo y cena, mientras que el resto de instituciones sólo reciben desayunos, siempre que sean parte de los siguientes segmentos: educación Inicial (entre 3 y 4 años), los alumnos de Educación General Básica (entre 5 y 14 años), y alumnos de bachillerato en jornada vespertina. El contenido nutricional ofrecido por el PAE contempla las diferentes necesidades de los alumnos según grupo de edad.

Tabla 9. Comparación de los programas de alimentación escolar en cinco países de América del Sur

a) Características generales

Variables	Perú: Qaliwarma		Chile: PAE	Colombia	Bolivia	Ecuador
	Actual	Propuesto				
Cobertura	Universal	Universal	Focalizado	Universal	Universal	Universal
Priorización	Modalidad de atención diferenciada según condición socioeconómica	Modalidad de atención diferenciada según condición socioeconómica y nutricional	Todos los alumnos que se encuentren dentro del 60% de las familias más vulnerables o con mayor desventaja socioeconómica del país tomando como referencia el Registro Social de Hogares.	<ul style="list-style-type: none"> • Zona Geográfica: Área rural y Área urbana. • Vulnerabilidad: Población de comunidades étnicas (indígenas, afrocolombianos, raizales, gitanos), población víctima del conflicto armado, población con vulnerabilidad nutricional y socioeconómica según los puntos de corte del SISBEN. 	Implementación gradual. <ol style="list-style-type: none"> 1. Municipios que no cuentan con programas de alimentación escolar y en los que cuentan además del desayuno escolar con servicio de almuerzo escolar. 2. Los municipios de mayor vulnerabilidad a la inseguridad alimentaria (102 municipios). 3. Municipios de vulnerabilidad media/baja (199 municipios) 4. Resto del país 	Según condición de la institución educativa: residencial o no residencial.
Modalidades de atención	Dos modalidades de atención: <ul style="list-style-type: none"> ▪ Desayuno y almuerzo ▪ Desayuno 	Tres modalidades de atención: <ul style="list-style-type: none"> ▪ Desayuno y almuerzo ▪ Desayuno ▪ Refrigerio 	Cinco modalidades de atención: <ul style="list-style-type: none"> • Desayunos • Almuerzos • Once (merienda) • Colaciones • Cenas 	Ocho posibles modalidades de atención: <ul style="list-style-type: none"> • Desayuno • Complemento alimentario jornada de la tarde • Almuerzo • Refrigerio¹ • Desayuno y refrigerio¹ • Complemento alimentario jornada de la tarde y refrigerio¹ • Desayuno y almuerzo¹ • Almuerzo y refrigerio¹ 	Cuatro modalidades de atención: <ul style="list-style-type: none"> • Desayuno • Desayuno y almuerzo • Desayuno /merienda o almuerzo • Solo almuerzo 	Según tipo de institución <ul style="list-style-type: none"> • En residencias estudiantiles (Tres modalidades): desayuno, almuerzo y cena. • Otras instituciones educativas: Desayunos (raciones alimenticias).

b) Criterios de asignación y priorización

Variables	Perú: Qaliwarma		Chile	Colombia	Bolivia	Ecuador
	Actual	Propuesto				
Criterio de asignación y a qué nivel	Socioeconómico a partir del mapa de pobreza a nivel distrital	Socioeconómico y nutricional a nivel de escuela	Focalización individual, diferenciado por estudiante dentro de las escuelas.	Socioeconómico a partir del SISBEN, ámbito de residencia y demás grupos vulnerables. Priorización en dos etapas, primero a nivel de escuelas y luego de forma individual según las características socioeconómicas de los alumnos.	El programa se prioriza según la clasificación de los municipios y se entrega a nivel de escuelas.	El programa es universal. La unidad de priorización es la escuela: se prioriza a escuelas residenciales para asignar las tres modalidades de atención. Al resto de escuelas sólo se le entrega un desayuno.
Contenido calórico “plano” o diferenciado	Contenido calórico “plano” diferenciado según nivel socioeconómico	Diferenciado según nivel socioeconómico y necesidades nutricionales.	Contenido diferenciado según las necesidades nutricionales del grupo de edad.	Contenido diferenciado según las necesidades nutricionales del grupo de edad.	Contenido nutricional de las raciones varía de un municipio a otro. Aunque no se ha estandarizado los requerimientos nutricionales para la población.	Contenido diferenciado según las necesidades nutricionales del grupo de edad.

1 Solo se ofrecerán dos complementos si es que se ha visto atendida el total de la población escolar en situación de vulnerabilidad financiada por el Ministerio de Educación de Colombia y los entes territoriales

Fuente: Junta Nacional de Auxilio Escolar y Becas (2019); Lavado, P., Lovón, M. y Vargas, R. (2018), Ministerio de Educación Nacional (2013); Ministerio de Educación de Bolivia (2015); Ministerio de Educación de Ecuador (2019); Proalimentos (2014). Elaboración propia

9. PROTOCOLO DE ACTUALIZACIÓN DEL ÍNDICE

La presente sección corresponde al protocolo de actualización de los índices de priorización elaborados en este estudio (en adelante, **Protocolo de Actualización**). La metodología que se implementó para la construcción de estos índices responde, en buena cuenta, a la información disponible a la que tuvo acceso el equipo consultor durante los plazos establecidos para la consultoría. Esta información no fue completa ni perfecta, pero sí lo suficientemente buena como para alcanzar los objetivos del estudio. En el caso del índice socioeconómico, se utilizó un índice de riqueza construido a partir de un análisis de componentes principales que, si bien se encuentra altamente correlacionado con el gasto y, en ese sentido, constituye una buena aproximación al nivel de pobreza monetaria de los alumnos, no es la misma variable (información imperfecta). En el caso del índice nutricional, se utilizó la prevalencia de sobrepeso/obesidad de los alumnos en las escuelas que, si bien es la mejor información con la que se cuenta para caracterizar el estado nutricional de los alumnos, se tuvo que imputar información al 40% de las escuelas (información incompleta).

9.1. SUPUESTOS

El presente Protocolo de Actualización ha sido diseñado considerando una frecuencia de actualización quinquenal y partiendo del supuesto de que se contará con información adecuada y completa. Caso contrario, no será posible actualizar ambos índices, al menos, con la frecuencia mencionada. Asimismo, sugerimos realizar la actualización en el segundo semestre del año escolar. De este modo, este Protocolo de Actualización parte de dos grandes supuestos:

i) El Estado contará con un Mapa de Pobreza actualizado a partir del Censo 2017¹¹.

A la fecha, el Perú cuenta con dos mapas de pobreza: uno publicado en 2009 y otro en 2013, ambos elaborados a partir de la información de los Censos Nacionales 2007 y con inferencia a nivel distrital. No obstante, se espera que en el corto plazo (próximo año) se cuente con un nuevo Mapa de Pobreza elaborado a partir de los últimos censos nacionales de población y vivienda del 2017. Asimismo, se tiene la expectativa que dicho Mapa de Pobreza permita tener una inferencia a mayor que el distrito (por ejemplo, localidades, barrios, o incluso manzanas). Las experiencias recientes del Instituto Nacional de Estadística e Informática – INEI¹², sugieren que el nuevo Mapa de Pobreza pretende ofrecer información más desagregada para la focalización de políticas públicas. En caso no se cuente con un Mapa de Pobreza actualizado, no será posible actualizar el índice socioeconómico debido a que la información que se utilizó para su construcción corresponde a la del Censo 2017. En el Perú, los censos nacionales se realizan con una frecuencia no menor a los diez años. El último Censo Nacional se llevó

¹¹ Los mapas de pobreza tienen como objetivo identificar las localidades con mayor incidencia de pobreza y, en ese sentido, constituye el instrumento más preciso con el que se cuenta para la focalización de las políticas públicas.

¹² En 2009 y 2013, el INEI publicó planos estratificados de Lima Metropolitana a nivel de manzana según ingreso per cápita del hogar y según grupos de pobreza monetaria.

a cabo hace dos años, por lo que no se espera contar con un nuevo censo en el corto o mediano plazo.

- ii) **El Estado continuará realizando mediciones de talla y peso a todos los alumnos matriculados en instituciones públicas con cierta periodicidad** (o, por lo menos, en las escuelas del ámbito de intervención del PNAEQW).

Para la construcción del índice nutricional se utilizó la información del PSE; no obstante, para su actualización no es necesario montar una evaluación de tal magnitud¹³, bastaría con contar con la talla y peso de todos los alumnos. En este estudio hemos validado que la información de talla y peso de hace tres años sirve para caracterizar el estado nutricional actual de los alumnos, toda vez que ambos índices se encuentran altamente correlacionados. No obstante, es evidente que con el pasar de los años, esta correlación será cada vez menor y que la información pasada ya no servirá para caracterizar el estado nutricional presente y, por tanto, será necesario contar con información actualizada. Consideramos que esto ocurrirá o bien porque el coeficiente de correlación entre ambos índices se volverá igual o menor a 0.5, o bien porque el parámetro de los modelos estimados ya no será significativo, lo que ocurra primera (ver **Anexo 12.2**). De acuerdo con nuestras estimaciones esto no ocurrirá antes de los cinco años; no obstante, el año exacto deberá validarse con un nuevo levantamiento de información. En caso no se disponga de la información censal sobre la talla y peso de los estudiantes, no será posible actualizar el índice nutricional dado que no existe ninguna otra fuente de información que podría servir como reemplazo.

9.2. BASES DE DATOS NECESARIAS

Dado lo anterior, la actualización de los índices de priorización requiere de la siguiente información:

- i) **Padrón de PNAEQW:** Corresponde al padrón de todas las instituciones educativas públicas de primaria que son atendidas por el PNAEQW a nivel nacional. Esta fuente de información determina el universo del estudio. Para la actualización de los índices de priorización es necesario que la base de datos contenga la siguiente información: código modular; código del local; número de anexo; localización geográfica; coordenadas geográficas; cantidad de alumnos atendidos por el programa; modalidad de atención y tipo de servicio que recibe la escuela; comité, ítem y región alimentaria al que pertenece la escuela; y, finalmente, si los alumnos de la escuela están considerados como población especial. La variable de identificación de esta base de datos debe ser el código modular del nivel de primaria de la institución educativa. Esta información tiene una frecuencia anual y puede ser solicitada a inicios de cada año a la Dirección Ejecutiva del programa. La institución responsable de proveer esta información debe ser el Ministerio

¹³ PSE es una evaluación integral de salud que, además del control de la talla y el peso, incluye tamizaje de calendario de inmunizaciones, tamizaje de hemoglobina, tamizaje de agudeza visual y detección de síntomas respiratorios.

de Desarrollo e Inclusión Social – MIDIS por ser el ente rector de los programas sociales a su cargo.

- ii) Padrón del SIAGIE:** Corresponde al padrón de todos los alumnos matriculados en primaria en las instituciones educativas públicas a nivel nacional (educación básica regular). Para la actualización del índice es necesario que la base de datos se encuentre nominada e incluya la siguiente información: DNI de los alumnos; nombres y apellidos; sexo; fecha de nacimiento; lugar de residencia; código modular y número de anexo de la institución educativa a la que pertenece el alumno; y nivel de atención de la institución educativa. La variable de identificación de esta base de datos debe ser el DNI de los alumnos. Esta información tiene una frecuencia anual y puede ser solicitada a partir del tercer trimestre del año escolar¹⁴. La institución responsable de proveer esta información debe ser el Ministerio de Educación (MINEDU) por ser la institución encargada del SIAGIE.

- iii) Censo 2017:** Corresponde al padrón de todas las personas censadas en el Perú y contiene todas sus características socioeconómicas. Independientemente de que no proporcione información más actualizada de la que ya se tiene, esta fuente de información contiene el padrón poblacional, el cual será de utilidad para identificar a los alumnos. Para la actualización del índice es necesario que la base de datos se encuentre nominada e incluya la siguiente información: nombres y apellidos de las personas; DNI; sexo; fecha de nacimiento; lugar de residencia; y todas las variables referidas a la localización de la vivienda y el número de hogares (sección I de la Cédula Censal del Censo 2017), características y servicios de la vivienda (sección II), características del hogar (sección III), personas que conforman el hogar (sección IV) y características de la población (sección V). La variable de identificación de esta base de datos debe ser el DNI de las personas.

- iv) Mapa de Pobreza:** Corresponde a los planos estratificados del Perú según condición de pobreza. Para la actualización del índice socioeconómico es necesario que la base de datos contenga la siguiente información: código de ubicación geográfica (a nivel de manzana o localidad, dependiendo del nivel de desagregación); coordenadas geográficas; ingresos o gastos promedio per cápita de los hogares de cada unidad geográfica; y clasificación de la unidad geográfica según condición de pobreza monetaria. Esta información podría tener una frecuencia quinquenal, dependiendo de los esfuerzos del INEI. La institución responsable de proveer esta información debe ser el INEI por ser la institución encargada del Mapa de Pobreza.

- v) Mediciones de talla y peso de los alumnos:** Corresponde a los resultados de la medición de talla y peso de los alumnos de las instituciones educativas públicas de primaria a nivel nacional. Para la actualización del índice es necesario que la base de datos se encuentre nominada e incluya la siguiente información: DNI; nombres y apellidos de los alumnos; sexo; fecha de nacimiento; lugar de residencia; fecha de medición; y la talla (en centímetros) y el peso (en kilogramos) de los alumnos. La variable de identificación de esta base de datos debe ser el DNI de los alumnos. La frecuencia de esta información

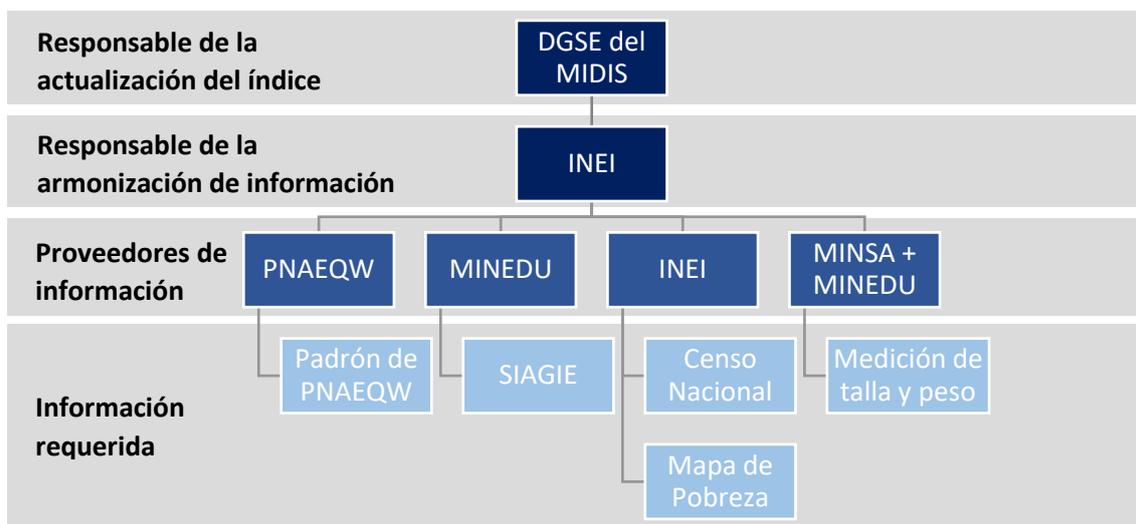
¹⁴ El Padrón del SIAGIE se cierra, en estricto, luego del tercer trimestre del año escolar.

debería ser de quinquenal. Las instituciones responsables de proveer esta información deberían ser el Ministerio de Salud (MINSA) junto con el MINEDU.

9.3. ACTORES RELEVANTES

De este modo, se tiene que las instituciones responsables de proveer la información necesaria para la actualización de los índices de priorización deberían ser el MIDIS (Dirección Ejecutiva del PNAEQW), el MINEDU, el INEI y el MINSA. Asimismo, el MIDIS, en su calidad de ente rector de la política social en el Perú, debería ser la institución responsable de la actualización de los índices de priorización. El órgano encargado de ello debería ser la Dirección General de Seguimiento y Evaluación (DGSE) del MIDIS. No obstante, debe apoyarse en el INEI para armonizar las bases de datos necesarias para su cálculo toda vez que se trata de información nominada.

Gráfico 8. Actores relevantes para la actualización de los índices de priorización



Elaboración propia.

9.4. PROCEDIMIENTO DE ACTUALIZACIÓN

La metodología de actualización de los índices de priorización consta de cuatro pasos: la armonización de las bases de datos, la construcción de índices a partir de información disponible, la imputación de datos faltantes, y la clasificación de escuelas en grupos relevantes. A continuación, el detalle de cada paso:

– Paso 1: Armonizar las bases de datos

La actualización de los índices de priorización requiere de la armonización de las cinco bases de datos ya mencionadas: el padrón del PNAEQW, el padrón del SIAGIE, el Censo

2017, el Mapa de Pobreza y el PSE. La institución responsable de ello debe ser el INEI. La armonización de las bases de datos se debe realizar del siguiente modo:

- **Primero, empatar el padrón del PNAEQW con el del SIAGIE.** El empate entre las dos bases de datos se debe realizar a partir del código modular de la institución educativa. Cabe precisar que ambas bases de datos deben ser del mismo año. Por ejemplo, si se utiliza el padrón del PNAEQW del 2020, el padrón del SIAGIE también debe ser del 2020. Al cruzar estas fuentes de información, se pueden dar tres resultados: **empate perfecto**, es decir, la escuela se encuentra en ambas bases de datos; **solo PNAEQW**, es decir, la escuela se encuentra en el padrón del PNAEQW, pero no en el SIAGIE; y **solo SIAGIE**, es decir, la escuela se encuentra en el padrón del SIAGIE, pero no en el PNAEQW. Para efectos de la actualización del índice, solo son relevantes los primeros dos resultados. Esto porque, el último resultado (solo SIAGIE) corresponde a escuelas que no son atendidas por el programa y, en ese sentido, deben ser descartadas. A continuación, se debe revisar las escuelas que aparecen en el segundo resultado (solo PNAEQW). En el caso de este estudio, la mayoría de los códigos modulares del PNAEQW que no pudieron ser empatados con el SIAGIE correspondían a escuelas que atendían bajo la modalidad de educación básica especial. Estas escuelas fueron incorporadas dentro del universo del estudio a solicitud de la DGSE del MIDIS; no obstante, para futuras actualizaciones, debe evaluarse si se las incorpora o se les da un tratamiento especial. En caso se incorporen al universo de escuelas, existen dos opciones: o se las trata como escuelas con información perdida (como se realizó en este estudio), o se amplía la solicitud de información del SIAGIE a todas las modalidades de educación a fin de recuperar esta información. El resultado del empate de estas bases de datos será la base principal para la actualización del índice ya que incluye al universo de escuelas.
- **Segundo, empatar la base de datos anterior con el Censo 2017.** El empate entre las dos bases de datos se debe realizar a partir del DNI del alumno. Del mismo modo, se pueden dar tres resultados: **empate perfecto**, es decir, el alumno se encuentra en ambas bases de datos; **solo SIAGIE**, es decir, el alumno se encuentra en el padrón del SIAGIE, pero no en el Censo 2017; y **solo Censo 2017**, es decir, la persona solo se encuentra en el Censo 2017. Del mismo modo que en el caso anterior, se debe descartar las observaciones que corresponden al tercer resultado y se debe revisar las que corresponden al segundo. Más adelante, estas últimas observaciones podrían necesitar algún tipo de imputación.
- **Tercero, empatar la base de datos anterior con el Mapa de Pobreza.** El empate entre la base de datos armonizada hasta el momento con el Mapa de Pobreza debe realizarse a partir de las coordenadas geográficas o el código de identificación de la manzana o el nivel de agregación inmediatamente superior que permita el mapa donde residen los alumnos (en adelante, **lugar de residencia**). Del mismo modo, se pueden dar tres resultados: **empate perfecto**, es decir, la localidad se encuentra en ambas bases de datos; **solo Censo 2017**, es decir, la localidad donde se encuentra la vivienda está en el Censo 2017, pero no se identifica en el Mapa de Pobreza; y, **solo Mapa de Pobreza**, es decir, la localidad donde se encuentra la vivienda está solo en el Mapa de Pobreza y no en el Censo 2017. De nuevo, se debe descartar todas las observaciones que corresponden al tercer resultado y se debe revisar las que

corresponden al segundo ya que este resultado podría deberse a un error de codificación. Si no se encuentra la manzana o la localidad de la vivienda en el Mapa de Pobreza, se debe subir al nivel de agregación inmediatamente superior e imputar dicho valor.

- **Cuarto, empatar la base de datos anterior con las mediciones de talla y peso de los alumnos.** El empate entre la base de datos armonizada con las mediciones de talla y peso de los alumnos a nivel censal debe realizarse a partir del DNI del alumno. Del mismo modo, se pueden dar tres resultados: **empate perfecto**, es decir, el alumno se encuentra en ambas bases de datos; **solo SIAGIE**, es decir, el alumno se encuentra en el padrón del SIAGIE, pero no en la base de datos de las mediciones de talla y peso; y **solo en la base de datos de las mediciones de talla y peso**, es decir, el alumno solo se encuentra en esta base. De nuevo, se deben descartar todas las observaciones que corresponden al tercer resultado y revisar las que corresponden al segundo. Más adelante, estas últimas observaciones podrían necesitar algún tipo de imputación.
 - **Finalmente, el resultado de este paso debe ser una base de datos armonizada a nivel de alumnos** que contenga las características socioeconómicas de los alumnos (incluyendo la condición de pobreza del lugar donde residen), la talla y peso de los alumnos, e información del servicio alimentario que ofrece PNAEQW.
- **Paso 2: Construcción de índices a partir de información disponible**

A continuación, se deben construir los índices de priorización sobre la base de datos armonizada y agregar la información a nivel de escuelas.

i) Índice socioeconómico:

El índice socioeconómico realizado para este estudio corresponde al puntaje del primer componente de un PCA. Para futuras actualizaciones, este índice deberá ser reemplazado por el gasto per cápita promedio del lugar de residencia de los alumnos. De este modo, se tendrá una base de datos a nivel de alumnos con sus respectivos niveles de gasto per cápita. Para agregar esta información a nivel de escuelas, debe sumarse todos los gastos per cápita promedio de los lugares donde residen los alumnos y dividirse entre el total de alumnos de cada escuela, de este modo, el índice socioeconómico estará dado por el gasto promedio per cápita promedio de los alumnos de la escuela, y tomará valores entre 0 y 1.

ii) Índice nutricional:

El índice nutricional está dado por la proporción de alumnos con sobrepeso u obesidad en cada escuela. Para estimar esta proporción, primero, se debe calcular el IMC de los niños a partir de la información de talla y peso del PSE. A continuación, se debe computar sus z-scores utilizando los patrones de referencia de la Organización Mundial de la Salud (OMS) para niños de 5 a 19 años. Luego, dichos z-scores deben ser comparados con los puntos de corte establecidos por la OMS para identificar el estado nutricional de los niños: delgadez ($<-2SD$), normal ($>=-2SD$ y

$\leq 1SD$), sobrepeso ($>1SD$ y $\leq 2SD$) y obesidad ($>2SD$). De este modo, se tendrá una base de datos que identifique el estado nutricional de cada niño. Esta información debe ser agregada a nivel de escuelas, de modo que sea posible identificar la cantidad de niños con sobrepeso y obesidad en cada escuela. Finalmente, el indicador se construye como la cantidad de alumnos con sobrepeso u obesidad en cada escuela como proporción de la cantidad total de alumnos en cada escuela, y tomará valores entre 0 y 1.

– **Paso 3: Imputación de datos faltantes**

Luego de armonizadas las bases de datos y antes de construir los índices de priorización, se evaluará las necesidades de imputación, pero este procedimiento solo se realizará una vez construidos ambos índices. En el caso del índice socioeconómico, no será necesario imputar información porque el procedimiento de armonización del Mapa de Pobreza contempla que, en caso no se encuentre el gasto per cápita promedio del lugar de residencia del alumno, se le imputará el dato del nivel inmediatamente superior. Por tanto, todos los alumnos contarán con información del gasto per cápita y, en consecuencia, lo mismo ocurrirá con las escuelas. En el caso del índice nutricional, sí podría ser necesario realizar algún procedimiento de imputación ya que podría darse el caso de que no todos los alumnos cuenten con información de talla y peso y, dependiendo de su concentración, podría darse el caso de que algunas escuelas no cuenten con la suficiente información como para inferir el estado nutricional de los niños. No obstante, consideramos que debe ser bastante menor por lo que sugerimos establecer el siguiente criterio para identificar si se requiere o no imputar información a una escuela. Para ello, primero, se debe calcular la proporción de alumnos dentro de cada escuela que tiene información completa de talla y peso. Segundo, se debe estimar el error muestral de cada escuela con la información disponible. Si una escuela tiene menos del 50% de alumnos con información completa de talla y peso y, además, tiene un error muestral superior al 10%, debe imputársele un dato. La metodología sugerida para realizar esta imputación es la del vecino más cercano (ver **Anexo 12.2**).

– **Paso 4: Clasificación de escuelas en grupos relevantes**

Como resultado del paso anterior, se cuenta con una base de datos a nivel de escuelas con sus respectivos índices con información completa (uno socioeconómico y otro nutricional). El último paso consiste en clasificar las escuelas en grupos relevantes. El objetivo es clasificar a las escuelas según el índice socioeconómico y el índice nutricional, para luego cruzar ambas clasificaciones y encontrar nueve categorías. En el caso del índice socioeconómico las escuelas deben ser clasificadas en grupos de alta, media y baja necesidad; y, en el caso del índice nutricional, en grupos de alto, medio y bajo sobrepeso/obesidad. A continuación, los puntos de corte para hacer dicha clasificación:

i) Índice socioeconómico:

Para la clasificación de escuelas a partir del índice socioeconómico, sugerimos utilizar un análisis de conglomerados por k-medias. El número de grupos debe establecerse en tres y los valores medios de cada grupo determinarán si corresponde a alta, media o baja necesidad. Cabe precisar que, como en el caso de

este estudio, el número óptimo de grupos puede que no sea tres, pero ello no invalida la metodología.

ii) Índice nutricional:

Se considera que una escuela tiene alto sobrepeso/obesidad, si el índice nutricional es superior a 0.2; tiene medio sobrepeso/obesidad, si el índice es menor de 0.2 pero mayor o igual a 0.05; y tiene bajo sobrepeso/obesidad, si el índice es menor a 0.05.

Finalmente, se deben cruzar ambas clasificaciones, obteniendo así 9 categorías.

10. CONCLUSIONES

Se ha construido una base de datos funcional a los objetivos del estudio. El objetivo del presente estudio fue elaborar y validar los índices de priorización propuestos en la Nota Metodológica para la prestación del servicio alimentario del PNAEQW en el nivel de educación de primaria. Para ello, primero se armonizaron cuatro bases de datos: el padrón de escuelas atendidas por el PNAEQW en el año 2019, el padrón de alumnos del SIAGIE 2017, el padrón de población y vivienda de los Censos Nacionales del 2017, y el Plan de Salud Escolar 2015-2016. De este ejercicio, se logró recuperar la información socioeconómica del 72% alumnos (en el 98% de escuelas del padrón) e información nutricional del 42% de alumnos (en el 91% de escuelas del padrón).

Se ha construido índices de priorización válidos a partir de la base de datos armonizada. La metodología propuesta en la Nota Metodológica fue revisada y adaptada para construir los índices de priorización (socioeconómico y nutricional), dada la información existente. En el caso del índice socioeconómico, se implementaron dos metodologías dando origen a dos índices distintos, uno construido a partir de la tasa de pobreza según el método de las NBI y otro a partir de un índice de riqueza construido a partir de un ejercicio de PCA. En el caso del índice nutricional, se utilizó un único indicador que ordena a las escuelas en función a la proporción de alumnos con sobrepeso u obesidad. A continuación, se realizó la imputación de la información faltante. Las estrategias implementadas permitieron recuperar al 6% de escuelas que no contaban con la suficiente información socioeconómica para caracterizar la condición de vulnerabilidad de los alumnos, y del 32% de escuelas que carecían de información nutricional o tenían información insuficiente.

Se ha categorizado a las escuelas en función a las dimensiones socioeconómica y nutricional, basados en los índices construidos. Con la información completa de todas las escuelas, el objetivo fue construir tres grupos en cada dimensión: alta, media y baja necesidad a partir del índice socioeconómico; alto, medio y bajo sobrepeso/obesidad a partir del índice nutricional. En el caso del índice socioeconómico, se realizaron tres ejercicios de agrupamiento (dos aproximaciones utilizando quintiles de riqueza y un análisis por conglomerados), mientras que en el caso del índice nutricional solo uno (a partir de umbrales definidos por el equipo técnico del PMA).

El índice socioeconómico basado en componentes principales y categorizado bajo la metodología de conglomerados es el que tiene mejor comportamiento. Si bien presentamos varias alternativas para realizar el agrupamiento de escuelas a partir del índice socioeconómico, sugerimos utilizar el índice de riqueza construido a partir del PCA, toda vez que se encuentra altamente correlacionado con el gasto del hogar y, en ese sentido, predice mejor el nivel de riqueza de los alumnos de las escuelas. Asimismo, sugerimos utilizar el agrupamiento por análisis de conglomerados (k-medias), no solo porque reduce la discrecionalidad en el agrupamiento de las escuelas, sino porque ofrece un agrupamiento que maximiza la homogeneidad de las escuelas al interior de cada grupo y la heterogeneidad de estos al exterior.

Se ha generado una tabla de priorización que define nueve categorías basadas en las dimensiones socioeconómicas y nutricionales. Finalmente, se realizó el agrupamiento de las escuelas bajo cada índice y se integraron los resultados, dando lugar a una matriz con nueve categorías (ver **Tabla 10**). Esta matriz permitirá diferenciar el servicio alimentario que el PNAEQW ofrece en las escuelas considerando la carga socioeconómica y nutricional de las escuelas. Esta priorización del servicio no solo pasa por diferenciar el tipo de menú que el programa ofrece a cada escuela sino también las características de este (tipos de combinaciones, grupos de alimentos y proporciones) y el paquete educativo y comunicacional. Para este estudio se consideraron tres modalidades de atención (desayunos, desayunos + almuerzos, y refrigerios), las cuales fueron sugeridas por el PMA sobre la base de la experiencia internacional y la asistencia técnica que el Programa Mundial de Alimentos brinda a los programas de alimentación escolar en el mundo. De acuerdo con esta propuesta, las escuelas que presentan una prevalencia de sobrepeso u obesidad media y alta podrían recibir la misma modalidad de atención independientemente de la su condición socioeconómica, por lo que esta matriz podría resumirse en seis categorías que involucran tres modalidades de atención.

Tabla 10. Distribución de escuelas y de alumnos según índices de priorización*
(alternativa sugerida por el equipo consultor)

Clasificación según índice nutricional	Clasificación según índice socioeconómico			Total
	Necesidad alta	Necesidad media	Necesidad baja	
Cantidad y porcentaje de escuelas				
Sobrepeso/obesidad alta	1,641	2,232	3,101	6,974
	6%	8%	11%	24%
Sobrepeso/obesidad media	10,192	3,816	258	14,266
	35%	13%	1%	48%
Sobrepeso/obesidad baja	6,887	1,237	122	8,246
	23%	4%	0%	28%
Total	18,720	7,285	3,481	29,486
	63%	25%	12%	100%
Cantidad y porcentaje de alumnos				
Sobrepeso/obesidad alta	67,699	319,742	1,160,855	1,548,296
	3%	12%	45%	60%
Sobrepeso/obesidad media	383,265	363,162	81,811	828,238
	15%	14%	3%	32%
Sobrepeso/obesidad baja	172,723	33,590	7,223	213,536
	7%	1%	0%	8%
Total	623,687	716,494	1,249,889	2,590,070
	24%	28%	48%	100%

*Nota: Los porcentajes están expresados respecto al total de escuelas y de alumnos atendidos por el PNAEQW.

Fuente: Base de datos armonizada. Elaboración propia.

Las modalidades de atención sugeridas por el equipo técnico del PMA no representan cambios muy significativos en términos operativos ni presupuestales a lo ofrecido por PNAEQW. Esto debido a que actualmente PNAEQW entrega 1,87 millones de raciones de desayunos y 716 mil desayunos más almuerzos. Bajo la propuesta alterativa, deberá más bien entregar 1,57 millones

de raciones de refrigerios, 464 mil desayunos y 556 mil desayunos más almuerzos. Estas diferencias no parecen implicar cambios operativos muy disruptivos, y más bien parecen ser presupuestalmente favorables pues se reduce la cantidad de raciones de la modalidad costosa (desayuno + almuerzo) y se aumenta las raciones de las modalidades baratas (refrigerio y desayuno).

Se elaboró un protocolo de actualización de los índices de priorización. El protocolo de actualización fue elaborado considerando una frecuencia de actualización quinquenal y bajo el supuesto de que el Estado contará con un Mapa de Pobreza actualizado a partir del Censo Nacional y de que continuará realizando mediciones de talla y peso a todos los alumnos matriculados en instituciones públicas con cierta periodicidad. Caso contrario, no será posible actualizar ambos índices, al menos, con la frecuencia mencionada. La institución responsable de la actualización del índice debería ser el MIDIS, quien necesitará contar con el apoyo del INEI para la armonización de la información requerida.

Queda como trabajo pendiente que el PNAEQW valide los resultados del estudio. En este informe se presenta una serie de alternativas para la construcción del índice socioeconómico y para el agrupamiento de las escuelas. Se sugiere que el PNAEQW revise cada una de ellas y elija aquella que se adecúe mejor al funcionamiento del programa.

11. BIBLIOGRAFÍA

Achrens, A.; C. Hansen y M. Schaffer (2019), "lassopack: Model selection and prediction with regularized regression in Stata", IZA DP N° 12081

Alkire, S., U. Kanagaratnam y N. Suppa (2018), "The Global Multidimensional Poverty Index (MPI): 2018 Revision".

Black, R. et al (2013), "Maternal and child undernutrition and overweight in low-income and middle-income countries".

Elbers, C.; J. Lanjouw y P. Lanjouw (2002) *Microlevel estimation of welfare*. Mimeo.

Elbers, C.; J. Lanjouw y P. Lanjouw (2002) "Microlevel estimation of poverty and inequality". *Econometrica* 71(1), 355-364.

Feres, J. y X. Mancero (2001), "El método de las necesidades básicas insatisfechas (NBI) y sus aplicaciones en América Latina". CEPAL: División de Estadística y Proyecciones Económicas.

Filmer, D. y L. Pritchett (2001), "Estimating wealth effects without expenditure data-or tears: an application to educational enrollments in states of India". *Demography*, Volume 38-Number 1, February 2001: 115-132.

Hastie, T.; R. Tibshirani y J. Friedma (2008), *The elements of statistical learning*, Spring series in statistics

JLM Solutions (7 de octubre de 2019). *Programa de Alimentación Escolar*. Recuperado de <https://www.moodle.ec/programa-de-alimentacion-escolar/>

Junta Nacional de Auxilio Escolar y Becas (Junaeb) (7 de octubre de 2019). *Programa de Alimentación Escolar (PAE)*. Recuperado de <https://www.junaeb.cl/programa-de-alimentacion-escolar>

Lavado, P., Lovón, M. y Vargas, R. (2018). *Priorización del Programa Nacional de Alimentación Escolar Qali Warma: Nota metodológica*.

Lavado, P., M. Lovón y R. Vargas (2018), "Nota Metodológica para la priorización del Programa Nacional de Alimentación Escolar Qali Warma".

McKenzie, D. (2003), "Measuring inequality with asset indicators". *Journal of Population Economics* (2005) 18:220-260.

Ministerio de Educación (7 de octubre de 2019). Proyecto Intervención en la Alimentación Escolar. Recuperado de <https://educacion.gob.ec/programa-de-alimentacion-escolar/>

Ministerio de Educación Nacional. (2013). *Lineamientos técnico-administrativos y estándares del programa de alimentación escolar (PAE)*. Recuperado de https://www.mineducacion.gov.co/1621/articles-235135_archivo_pdf_lineamientos_tecnicos.pdf

Ministerio de Educación. (2015). *Programa Nacional De Alimentación Complementaria Escolar (PNACE 2015 - 2020)*. Recuperado de <http://extwprlegs1.fao.org/docs/pdf/bol152754.pdf>

Minujin, A. y JH Bang (2002), "Indicadores de inequidad social. Acerca del uso del índice de bienes para la distribución de los hogares". *Desarrollo Económico* 42(162): 129-146.

OMS (2009), "Who AnthroPlus for Personal Computers Manual. Software for assessing growth of the world's children and adolescents".

Proalimentos (2014). *Intervención en la alimentación escolar*. Recuperado de <https://www.finanzas.gob.ec/wp-content/uploads/downloads/2017/03/PROYECTO-DE-INTERVENCION-EN-LA-ALIMENTACION-ESCOLAR-15-1-2015.pdf>

Prudencio, Julio (2013). *Programa Nacional De Alimentación Complementaria Escolar: PNACE 2013 - 2020*. Recuperado de <http://extwprlegs1.fao.org/docs/pdf/bol152754.pdf>

Richman, M.; T. Trafalis e I. Adrianto (2009), "Missing data imputation through machine learning algorithms". En S. E. Haupt et al. (eds.), *Artificial Intelligence Methods in the Environmental Sciences*.

Trujillo, Yadira (7 de mayo de 2019). Focalización de alimento escolar se revisa. *El Comercio*. Recuperado de <https://www.elcomercio.com/actualidad/alimento-escolar-estudiantes-educacion-fiscal.html>.

12. ANEXOS

12.1. ANEXO METODOLÓGICO

En el presente anexo incluye el detalle de la metodología empleada para la construcción de los índices de priorización. Este anexo se organiza del siguiente modo. Primero se presenta la metodología empleada para la armonización de las bases de datos necesarias para la construcción de los índices. Luego, se presenta la metodología para la construcción del índice socioeconómico y del índice nutricional, y los resultados obtenidos. En la medida que la construcción ambos índices requiere de un procedimiento de imputación, primero se presenta la metodología para la construcción de los índices sobre la base de los datos disponibles, luego se presenta el proceso de imputación realizado para completar la data faltante, y finalmente se exponen los resultados de la clasificación realizada con la información completa.

1. ARMONIZACIÓN DE LAS BASES DE DATOS

A continuación, se presenta la descripción de las bases de datos empleadas para la construcción de los índices de priorización y el detalle el procedimiento de armonización de las bases de datos necesarias para la construcción de los índices de priorización.

1.1. Bases de datos empleadas

Como parte de las actividades realizadas, se armonizaron las siguientes bases de datos:

- **Padrón de PNAEQW 2019** (en adelante, **PNAEQW 2019**): La base de datos PNAEQW 2019 contiene la relación de escuelas públicas que son atendidas por el programa actualizado al 2019¹⁵. Dicho padrón incluye la siguiente información de cada escuela: localización geográfica, datos de identificación (código modular, código del local, número de anexo, dirección, coordenadas geográficas, área (urbano/rural) y nivel de atención), cantidad de alumnos atendidos por el programa, comité al que corresponde¹⁶, ítem al que

¹⁵ Ver Resolución de Dirección Ejecutiva N°451-2018-MIDIS/PNAEQW, 487-2018-MIDIS/PNAEQW, Resolución de Dirección Ejecutiva N°D000004-2019-MIDIS/MIDIS/PNAEQW-DE y Resolución de Dirección Ejecutiva N°D000011-2019-MIDIS/MIDIS/PNAEQW-DE.

¹⁶ **Comité:** Es una organización que congrega a los representantes de la sociedad civil y entidades públicas que tienen el interés y/o la competencia para apoyar el cumplimiento de los objetivos del PNAEQW. Cuentan con capacidad jurídica para la compra de bienes y otros actos establecidos en las directivas que apruebe el PNAEQW para la provisión del servicio alimentario.

corresponde¹⁷, modalidad de atención que recibe¹⁸, tipo de servicio que recibe¹⁹, región alimentaria a la que pertenece²⁰, y si incluye población especial²¹. El PNAEQW 2019 tiene registradas a 63,415 instituciones educativas que son atendidas por el programa (contabilizadas a partir de sus códigos modulares), de las cuales 32,840 corresponden a inicial (52%), 29,504 a primaria (46%) y 1,071 a secundaria (2%). La variable de identificación de escuelas en esta base de datos es el código modular. Para el presente estudio, solo se consideran a las instituciones públicas de nivel de primaria, las cuales albergan a alrededor de 2.6 millones de estudiantes (67% del total de usuarios de PNAEQW).

- **Base de datos del SIAGIE 2017:** Esta base de datos contiene el padrón de todos los alumnos matriculados en instituciones públicas y privadas a nivel nacional. La variable de identificación de alumnos en esta base de datos es el DNI. Para fines del presente estudio, se solicitó el padrón de las instituciones públicas de primaria del 2017, el cual asciende a 2,630,889 alumnos y 29,636 escuelas. Ese directorio constituye el universo del estudio.
- **Censos Nacionales 2017: XXII de Población, VII de Vivienda y III de Comunidades Indígenas** (en adelante, **Censo 2017**): El Censo 2017 contiene el padrón y las características socioeconómicas de todas las personas censadas en el último censo de población y vivienda realizado en el Perú en 2017 (el censo anterior corresponde a 2007). De acuerdo con el INEI, se logró censar al 91% de la población (29.4 de los 32.2 millones de personas que se estima que residen en el Perú), por lo que constituye la fuente de información más completa y actualizada para caracterizar a la población. El Censo 2017 contiene la siguiente información: (i) *localización de la vivienda y número de hogares*, (ii) *características y servicios de la vivienda* (tipo de vivienda, condición de ocupación, régimen de tenencia, materiales de construcción, y acceso a servicios básicos), (iii) *características del hogar* (combustible utilizado para cocinar, tenencia de activos, y migración), (iv) *personas que conforman el hogar*, y (v) *características de la población* (sexo, edad, estado civil, lengua materna, autoidentificación étnica, religión, fecundidad, discapacidad, educación y empleo). La variable de identificación de personas en esta base de datos es el DNI. Para fines del presente estudio, se solicitó la información de las viviendas y los hogares de todos los

¹⁷ **Ítem:** Es el conjunto de instituciones educativas públicas agrupadas por el PNAEQW para operacionalizar el proceso de compra para la provisión del servicio alimentario.

¹⁸ **Modalidad de atención:** Es el tipo de producto que se les entrega a las instituciones educativas. Existen dos modalidades de atención: raciones y productos. La modalidad de raciones consiste en la entrega de alimentos y bebidas listos para el consumo de los alumnos, mientras que la modalidad de productos consiste en la entrega de productos para que se preparen los desayunos y/o almuerzos en las instituciones educativas.

¹⁹ **Tipo de servicio:** PNAEQW ofrece dos tipos de servicio: desayunos y almuerzos.

²⁰ **Región alimentaria:** Es el resultado de la clasificación de las provincias del país de acuerdo a elementos culturales, geográficos, ecológicos, económicos y sociales en común. De este modo, se definen ocho regiones alimentarias por compartir una identidad culinaria en común.

²¹ **Poblaciones especiales:** Reciben tratamiento especial aquellas instituciones públicas ubicadas en los pueblos indígenas de la Amazonía peruana comprendidos en la base de datos oficial de pueblos indígenas listados en la RM N°321-2014-MC.

alumnos identificados en el SIAGIE 2017, así como de las características de estos y de todas las personas que conforman su hogar.

- **Plan de Salud Escolar** (en adelante, **PSE**): El PSE contiene los resultados de la evaluación integral de salud a los alumnos de los niveles educativos de inicial y primaria del ámbito de intervención del PNAEQW realizada en el marco del Plan de Salud Escolar 2013-2016 (ver Decreto Supremo N°010-2013-SA). Como parte de la evaluación integral, se realizaron controles de talla y peso, evaluación nutricional (IMC), tamizaje de calendario de inmunizaciones, tamizaje de agudeza visual, tamizaje de hemoglobina y detección de sintomáticos respiratorios a los estudiantes de las escuelas beneficiarias. Por tanto, la base del PSE contiene el registro de todos los alumnos evaluados y el resultado de todas las evaluaciones realizadas. La variable de identificación de los alumnos en esta base de datos es el DNI. Sin embargo, el PSE no constituye información censal pues las evaluaciones solo se realizaron a aproximadamente el 70% de los estudiantes atendidos por PNAEQW en cada año, por lo que será necesario imputar la información faltante. Para fines del presente estudio, se solicitaron los resultados de talla, peso e IMC de todos los alumnos de primaria evaluados en los años que se realizó el Plan de Salud Escolar, tanto de las variables continuas como de las categóricas. No obstante, solo se utilizó la información de los años 2015 y 2016 a fin de maximizar la comparabilidad de la información perdiendo la menor cantidad de datos posibles.

1.2. Procedimiento para la armonización de las bases de datos

La armonización de las bases de datos se realizó en tres etapas. Primero, se realizó el emparejamiento entre el PNAEQW 2019 y el SIAGIE 2017. Luego, se realizó el empate con el Censo 2017. Finalmente, se realizó se pegó esta base de datos con la del PSE²². La mayor parte del trabajo realizado estuvo a cargo del INEI y tuvo una duración total de 6 meses. A continuación, el detalle de la metodología seguida.

- **Empate PNAEQW 2019 – SIAGIE 2017**: El empate entre el directorio de PNAEQW 2019 y el SIAGIE 2017 se realizó a partir del código modular de las escuelas. De las 29,486 escuelas registradas en el directorio de PNAEQW 2019, se encontraron 29,087 en el SIAGIE 2017 (99%), ubicando así a 2.5 millones de alumnos (99%). De este modo, solo no se lograron empatar 399 códigos modulares con el SIAGIE. De acuerdo con el equipo técnico del MIDIS, los motivos de ellos son los siguientes: el 73% de los casos corresponden a escuelas que

²² Originalmente, también se buscó empatar la información del Censo 2017 con la ENAHO de los años 2013-2017 y la Encuesta Demográfica y de Salud Familiar (ENDES) de los años 2013 – 2017. Esto debido a la sugerencia de la Nota Metodológica de utilizar la información de la ENAHO para predecir el nivel de gasto de los hogares de los alumnos y la información de la ENDES para predecir su estado nutricional. Si bien se realizó la solicitud al INEI, el tiempo adicional que requería esta institución para realizar el empate de estas bases de datos sobrepasaba por mucho el cronograma de este estudio. Dentro de los plazos establecidos, el equipo consultor solo tuvo acceso al pegado del Censo 2017 con la ENAHO 2017 (9 mil observaciones), con lo cual se tuvo que descartar el uso del gasto per cápita. Del mismo modo, se descartó el uso de la ENDES.

intervienen bajo la modalidad de educación básica especial y no de educación regular considerada en este estudio; el 23% corresponde a escuelas de gestión privada que son atendidas por el PNAEQW; y el restante se debe a errores de codificación. Si bien sugerimos excluir dichas escuelas del presente estudio, estas fueron mantenidas a solicitud del equipo técnico del MIDIS. Como se verá más adelante, el tratamiento que recibieron estas escuelas fue el mismo que hubiera recibido una escuela que no hubiese contado con información socioeconómica y/o nutricional. No obstante, advertimos que existe más incertidumbre sobre la efectividad de este procedimiento en este subconjunto de escuelas debido a que no contamos con información apriorística para garantizar su comparabilidad. Más bien, presumimos que al ser escuelas que operan bajo modalidades diferentes de la básica regular, que son las que figuran en el SIAGIE, es más probable que sean diferentes. Sin embargo, no espera que su inclusión distorsione los resultados debido a que solo son 399. En adelante, denominaremos a la base de datos que surge de este empate como “base de datos principal” pues contiene el universo de alumnos y de escuelas del estudio.

- **Empate base de datos principal – Censo 2017:** El emparejamiento entre la base de datos principal y el Censo 2017 estuvo a cargo del INEI y se realizó a partir de la similitud en el registro de campos afines en ambas bases de datos, empleando criterios de aceptación de enlace determinístico y probabilístico. Los campos afines considerados para realizar el emparejamiento fueron los siguientes: DNI, nombres, apellido paterno, apellido materno, fecha de nacimiento (día, mes y año), sexo y código de ubigeo. Primero, se realizó el empate considerando el criterio de aceptación determinístico, es decir, si la coincidencia entre la información registrada en todos los campos afines mencionados era exacta (mismo DNI, mismos nombres y apellidos, mismo sexo, misma fecha de nacimiento y mismo código ubigeo). De este modo, se logró empatar al 58% de las observaciones del SIAGIE 2017. En el caso en que no se contara con el registro del DNI en alguna base de datos, bastó con que la coincidencia en el resto de los campos afines sí fuese exacta para considerarlo como un empate exitoso. Así, se empató al 30% de las observaciones del SIAGIE 2017. Luego, con el resto de las observaciones, se realizó un emparejamiento probabilístico, donde si bien la coincidencia entre los campos afines no es exacta, la similitud entre ellos es elevada, logrando emparejar de este modo al 9% de alumnos. Finalmente, se realizó una revisión manual del resto de observaciones que no se lograron pegar bajo ninguno de los dos métodos anteriores, logrando recuperar el 3% de observaciones del SIAGIE 2017.

Así, del total de alumnos emparejados anteriores (2.5 millones), se logró ubicar al 73% en el Censo 2017 (1.8 millones de alumnos), proporción que pasó a 72% luego de realizar la limpieza de la base de datos.

- **Empate base de datos principal – PSE:** El emparejamiento entre ambas bases de datos se realizó a partir del DNI del alumno y estuvo a cargo del INEI. Del total de alumnos de la base de datos principal (2.5 millones), se logró ubicar al 42% en el PSE (1 millón de alumnos).²³

²³ Al respecto, cabe precisar que el empate de la base de datos principal con la información del PSE se realizó en dos etapas. En la solicitud inicial para la armonización de las bases de datos nominadas, el MIDIS

El resumen del emparejamiento realizado se presenta en la **Tabla 11**. En adelante, denominaremos **base armonizada** al resultado del empate de las bases de datos del PNAEQW 2019, SIAGIE 2017, Censo 2017 y PSE, a nivel de alumnos. Estos resultados implican ciertas necesidades de imputación para recuperar la información perdida en el proceso de pegado. El procedimiento de imputación se realizó de manera diferenciada para el índice socioeconómico como para el nutricional, por lo que se detalla en la siguiente sección.

Tabla 11. Resultados del emparejamiento de bases de datos: PNAEQW 2019, SIAGIE 2017, Censo 2017 y PSE

	Alumnos		Escuelas	
	Nro.	%	Nro.	%
1) Empate PNAEQW 2019 – SIAGIE 2017				
Resultado del empate	2,497,346	99%	29,087	99%
Adicionales PNAEQW 2019	15,782(*)	1%	399	1%
Base principal (universo)	2,513,128	100%	29,486	100%
2) Empate base principal – Censo 2017				
Resultado del empate	1,822,988	73%	28,765	98%
Base de datos limpia	1,804,232	72%	28,742	98%
3) Empate base principal – PSE				
Resultado del empate	1,061,151	43%	26,843	92%
Base de datos limpia	1,038,754	42%	26,787	91%

(*) Aproximación sobre la base del número de usuarios.

Fuente: PNAEQW 2019, SIAGIE 2017, Censo 2017 y PSE. **Elaboración propia.**

solicitó empatar la base de datos principal con la base de datos del PSE con la que contaban. El INEI atendió su solicitud y devolvió una base de datos armonizada a nivel de alumnos, donde cada alumno ubicado en ambas bases de datos contaba con información de talla y peso. No obstante, la base de datos con la que se contaba presentaba una serie de inconsistencias que impedían medir el estado nutricional de los alumnos. Ante ello, el equipo consultor sostuvo una reunión con el equipo técnico del Seguro Integral de Salud (SIS) a fin de absolver las dudas que se tenían sobre la base de datos. Luego de revisar la información con la que se contaba hasta ese momento, el equipo del SIS confirmó la existencia de errores en la base de datos. Ante ello, sugirieron realizar una nueva solicitud de información para tener acceso a la data correcta que permitiera caracterizar el estado nutricional de los alumnos (talla y peso de todos los alumnos). El MIDIS procedió de ese modo y se le alcanzó una nueva base de datos. No obstante, en esta oportunidad, no fue posible solicitar al INEI que realice el empate con la base de datos principal debido a que el plazo adicional que requería la institución para realizar el pegado escapaba del cronograma del estudio. Ante ello, se procedió a realizar el pegado de manera indirecta y en dos etapas. Primero, se empató la segunda versión del PSE con la primera versión del PSE a partir de un identificador común en ambas bases de datos. Luego, se empató la información que pegó exitosamente con la del SIAGIE a fin de recuperar el identificador que vinculaba esta base con el resto de las bases de datos armonizadas por el INEI. Por este motivo, la tasa de éxito fue bastante menor a comparación de los otros ejercicios y solo se logró recuperar la información de talla y peso del 42% de los alumnos.

2. CONSTRUCCIÓN DE LOS ÍNDICES DE PRIORIZACIÓN

A continuación, se presenta el detalle de la metodología empleada para la construcción del índice socioeconómico y del índice nutricional.

2.1. Índice socioeconómico

De acuerdo con lo señalado en la Nota Metodológica, el índice socioeconómico se debía construir a partir del gasto per cápita o la tasa de pobreza monetaria de los hogares de los alumnos de las escuelas. Para ello, se sugería empatar la información de la Encuesta Nacional de Hogares (ENAH) 2013 – 2017 con la del Censo 2017, con lo cual sería posible modelar económicamente el gasto per cápita de los hogares en función de características socioeconómicas disponibles en el Censo, para luego predecir los valores correspondientes entre los hogares de alumnos con información faltante. Posteriormente, esto se colapsaría a nivel de las escuelas (o grupo de escuelas) de los distritos donde interviene el PNAEQW. No obstante, como ya se mencionó, la información de la ENAH a la que tuvo acceso el equipo consultor (año 2017) resultó insuficiente para realizar este ejercicio. Por ello, se exploraron otras opciones que emplearan variables no monetarias y que, además, hayan sido recogidas en el Censo 2017, que es la fuente de información con la que se cuenta para realizar este estudio. De esta manera, si bien no se modelaba y predecía el gasto per cápita, se buscaba un indicador los más correlacionado posible con esta variable.

Se exploraron tres alternativas. La primera fue medir el nivel de vulnerabilidad socioeconómica de los alumnos de las escuelas a partir de un enfoque de pobreza no monetaria que considere múltiples dimensiones que vayan más allá del nivel de ingresos o gastos de los hogares. Si bien existen múltiples metodologías para hacerlo, una de las más extendidas es la desarrollada por *Oxford Poverty and Human Development Initiative* (OPHI) en 2010 para el cálculo del Índice de Pobreza Multidimensional (**MPI**, por sus siglas en inglés). El MPI identifica las múltiples privaciones que enfrentan los hogares e individuos en términos de condiciones de vida, educación y salud, y las agrega en un índice que refleja tanto la incidencia como la intensidad de la pobreza de las personas. Actualmente, el MPI es considerado una medida estandarizada de pobreza multidimensional y se calcula de manera anual para de 100 países. El Perú no ha sido ajeno a esta tendencia y también han existido esfuerzos desde el MIDIS para incorporar el MPI como una de las medidas oficiales de pobreza. Por estos motivos, en un inicio, consideramos elaborar un MPI para la construcción del índice socioeconómico. No obstante, formalmente, el MPI requiere que todos los indicadores necesarios para su cálculo provengan de la misma base de datos (o, si provienen de distintas bases de datos, que puedan ser armonizadas a nivel individual) y, en el Perú, no contamos con una base de datos a nivel censal que recoja los indicadores de salud que se requieren para la construcción del MPI (en concreto, estado nutricional de los miembros del hogar y mortalidad infantil), por lo que no sería posible medir una de las dimensiones de este indicador (salud).

Ante la imposibilidad de implementar un MPI con los datos disponibles para el nivel de desagregación esperado (escuelas), se recurrió a una segunda alternativa basada en la metodología más conocida y utilizada en América Latina, y en el Perú, para la medición de pobreza multidimensional: el método de las NBI. Este método fue introducido por la Comisión Económica para América Latina y el Caribe (CEPAL) en los años ochenta, y tiene la ventaja de que se construye sobre la base de la información que se recoge en la mayoría de los censos y las encuestas de hogares de la región. No obstante, en la medida que es una metodología funcional a la información disponible, descansa en un marco teórico de limitado desarrollo, y mide apenas unas pocas necesidades específicas, dejando de lado otros elementos relevantes del bienestar. Asimismo, tiene la desventaja de que los indicadores utilizados y los niveles críticos de satisfacción – o al menos aquellos que se utilizan en el Perú – no se actualizan con frecuencia (en el Perú, no se actualizan desde 1993), por lo que podrían no estar caracterizando adecuadamente un estado general de privación y, en ese sentido, tener una limitada representatividad del estado de vulnerabilidad de la población. Aun así, constituye una de las metodologías más sencillas de replicar, explicar y actualizar, por lo que fue una de las seleccionadas en este estudio.

En contraste con el enfoque de la pobreza no monetaria, la tercera alternativa para la construcción del índice socioeconómico parte de un enfoque más pragmático donde se utiliza el análisis de componentes principales para construir un índice de riqueza que se encuentre altamente correlacionado con el ingreso y gasto de los hogares. Esta metodología fue propuesta por Filmer y Pritchett (2001) y ha sido ampliamente utilizada para aproximar el nivel de riqueza de los hogares en ausencia de información monetaria. En ese sentido, consideramos que constituye un buen reemplazo a la propuesta de la Nota Metodológica por lo que es la segunda alternativa que ofrece este estudio.

2.1.1. Construcción del índice sobre los datos disponibles

A continuación, se presenta el detalle de las dos metodologías implementadas para la construcción del índice socioeconómico y los resultados obtenidos.

2.1.1.1. Método de las Necesidades Básicas Insatisfechas

Como ya se mencionó, la primera alternativa para la construcción del índice socioeconómico fue utilizar el método de las NBI. Este método identifica a un hogar como pobre a partir de una serie de indicadores que miden ciertas dimensiones de la pobreza que van más allá de lo estrictamente monetario. En el Perú, el cálculo oficial de las NBI está a cargo del INEI y se realiza de manera anual a partir de la ENAHO, o cuando se cuenta con información censal actualizada. El INEI define cinco NBI de acuerdo al siguiente detalle:

1. **NBI 1 (vivienda inadecuada):** Población que reside en viviendas con características físicas inadecuadas, es decir en: (a) viviendas con paredes exteriores de estera; (b) viviendas con piso de tierra y paredes exteriores de quincha, piedra con barro, madera

u otros materiales; o, (c) viviendas improvisadas (cartón, lata, ladrillos, adobes superpuestos, entre otros).

2. **NBI 2 (hacinamiento):** Población que reside en viviendas con más de 3.4 personas por habitación, sin contar baños, cocina, pasadizo ni garaje.
3. **NBI 3 (vivienda sin servicios higiénicos):** Población que reside en viviendas que no disponen de servicio higiénico, es decir, que su servicio higiénico está conectado a: (a) río, acequia, canal o similar; o, (b) a campo abierto o al aire libre.
4. **NBI 4 (inasistencia escolar):** Población que pertenece a hogares donde al menos un niño de 6 a 12 años pariente del jefe del hogar no asiste a la escuela.
5. **NBI 5 (dependencia económica):** Población que pertenece a hogares con alta dependencia económica, es decir, aquella que reside en hogares cuyo jefe de hogar tiene primaria incompleta (hasta segundo año de primaria o tercer grado de primaria) y donde, además, existen más de tres dependientes por persona ocupada en el hogar (en caso ningún miembro se encuentre ocupado, se considera a los hogares con más de tres miembros en el hogar).

Estos indicadores se calculan para cada uno de los hogares y corresponden a variables dicotómicas que toman valor de 1 si la persona presenta la necesidad y 0 si no lo hace. Para identificar si una persona es pobre, primero se suman todas las NBI que presenta una persona del siguiente modo:

$$inbi_i = nbi1_i + nbi2_i + nbi3_i + nbi4_i + nbi5_i$$

donde: $inbi_i$ es un índice de recuento que indica cuantas NBI tiene el individuo i , $nbi1_i$ indica si el individuo i reside en una vivienda inadecuada (toma el valor de 1 si lo hace y de 0 en caso contrario); $nbi2_i$ indica si el individuo i reside en una vivienda con hacinamiento (toma el valor de 1 si lo hace y de 0 en caso contrario); $nbi3_i$ indica si el individuo i reside en una vivienda sin servicios higiénicos (toma el valor de 1 si lo hace y de 0 en caso contrario), $nbi4_i$ indica si el individuo i pertenece a un hogar donde algún niño no asiste a la escuela (toma el valor de 1 si lo hace y de 0 en caso contrario), y $nbi5_i$ indica si el individuo i pertenece a un hogar con alta dependencia económica (toma el valor de 1 si lo hace y de 0 en caso contrario). Luego, se define como pobres a todas aquellas que personas que presentan al menos una NBI.

Si $inbi_i = 0$, el individuo no es pobre.

Si $inbi_i \geq 1$, el individuo sí es pobre.

La tasa de pobreza se calcula como la sumatoria de personas identificadas como pobres dividida entre el total de la población:

$$pobreza_{nbi} = \frac{\sum_{i=1}^n inbi_i}{n}$$

donde n es la población total.

Para este estudio, se realizaron algunas adaptaciones al procedimiento anterior. A diferencia de la metodología descrita, la unidad de análisis en este caso es el alumno, por lo que las NBI se calcularon para cada uno de los alumnos de las escuelas. Luego, una vez identificados a los alumnos pobres, se agregó la información a nivel de escuelas, obteniendo así la proporción de alumnos dentro de cada escuela que se encuentra en situación de pobreza no monetaria. El índice socioeconómico está dado por esta variable (tasa de pobreza según NBI de las escuelas) y toma valores entre 0 y 1.

En la **Tabla 12** se presentan los resultados del cálculo de las NBI para la población de estudio. En la columna (1) se presentan los indicadores a nivel de alumnos (antes de agregar el indicador a escuelas); en la columna (2) se presenta el valor promedio de los indicadores luego de ser agregados a nivel de escuelas; y, finalmente –y únicamente con fines comparativos–, en la columna (3) se presentan los indicadores a nivel de alumnos calculados a partir de la ENAHO para la misma población de estudio, es decir, para los alumnos de instituciones públicas de primaria. No se presentan los indicadores agregados a nivel de escuelas a partir de la ENAHO porque esta fuente de información no permite identificar escuelas. Además, para cada caso, se presenta el dato para el total, el área urbana y el área rural. Como se observa, el 40% de los alumnos del PNAEQW se encuentra en situación de pobreza según NBI a utilizando la información del Censo 2017, resultado 8pp mayor al encontrado en la ENAHO 2017 para la misma población. Al agregar este resultado a nivel de escuelas se tiene que, en promedio, la mitad de los estudiantes en cada escuela se encuentra en situación de pobreza.

Tabla 12. Indicadores de NBI a nivel de alumnos

	PNAEQW - Alumnos (1)			PNAEQW – Escuelas (2)			ENAHO 2017 - Alumnos (3)		
	T	U	R	T	U	R	T	U	R
Con al menos una NBI	40%	33%	56%	55%	39%	61%	32%	25%	44%
NBI 1: Vivienda inadecuada	15%	13%	19%	19%	14%	21%	11%	9%	13%
NBI 2: Hacinamiento	21%	17%	29%	29%	20%	32%	15%	12%	19%
NBI 3: Vivienda sin SS.HH.	9%	4%	20%	20%	9%	24%	12%	6%	22%
NBI 4: Inasistencia escolar	4%	3%	7%	7%	4%	8%	2%	1%	2%
NBI 5: Dependencia económica	5%	3%	10%	8%	4%	9%	3%	2%	4%

(1) Calculado sobre la base del total de alumnos (1,804,232 alumnos). (2) Calculado sobre la base del total de escuelas con información suficiente (27,552 escuelas). (3) Calculado sobre la base del total de alumnos de instituciones públicas de primaria identificados en la ENAHO 2017 (11,020 alumnos).

Fuente: Base de datos armonizada y ENAHO 2017. **Elaboración propia.**

2.1.1.2. Análisis de Componentes Principales

La segunda alternativa seleccionada fue utilizar el componente principal de un PCA construido sobre la base de las características de las viviendas, el acceso a servicios básicos y la tenencia de activos durables en el hogar (en adelante, **activos**), como *proxy* del nivel de riqueza de los alumnos de las escuelas. Debido a su alta correlación con el nivel de ingreso y/o gasto de los hogares, esta metodología ha sido bastante favorecida por la literatura especializada para aproximar la riqueza de los hogares en ausencia de variables monetarias (ver Filmer y Pritchett (2001), Minujin y Bang (2002), McKenzie (2003), entre otros).

El PCA es una técnica estadística que permite reducir la dimensionalidad de un grupo de datos, encontrando combinaciones lineales no correlacionadas de las variables con la mayor varianza. Dentro de estas combinaciones, el primer componente es la combinación lineal que captura la mayor variabilidad de los datos a comparación del resto de componentes y, por tanto, el que utilizaremos para construir el índice socioeconómico. Como mencionan Filmer y Pritchett (2001), el supuesto detrás de ello es que la riqueza de los hogares (o su nivel de vida) explica la mayor varianza en los activos que poseen los hogares.

De manera formal se tiene que, dado un vector de activos x , el primer componente principal de las observaciones (y) está dado por la combinación lineal:

$$y = \alpha_1 x_1 + \alpha_2 x_2 + \dots + \alpha_N x_N$$

cuya varianza es mayor entre el resto de las combinaciones lineales, sujeto a la restricción $\alpha' \alpha = 1$, donde $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_N)'$ es el vector de coeficientes. Así, el puntaje del primer componente principal para el alumno i , dado el vector de activos x_i , corresponde a $y_i = \alpha' x_i$. El puntaje que se obtiene del primer componente principal constituye una suerte de índice que proporciona la mayor discriminación entre los alumnos, donde los activos que presentan una mayor varianza serán los que tendrán un mayor peso. Por ejemplo, un activo con el que los hogares de todos los alumnos cuentan tendrá un peso igual a cero en el primer componente principal, toda vez que no explica ninguna diferencia entre los alumnos, y lo mismo ocurrirá con un activo con el que ningún alumno cuente en su hogar. Por el contrario, un activo cuya distribución sea más desigual entre los alumnos tendrá un mayor peso en el primer componente. De este modo, el puntaje calculado (y_i) será una aproximación al nivel de riqueza del hogar de cada alumno. En este punto cabe mencionar que el puntaje del primer componente podría tomar valores menores a 0. No obstante, y como se verá más adelante, esto no tiene ninguna implicancia en el índice socioeconómico toda vez que lo único que requerimos es que exista un ordenamiento entre los alumnos de acuerdo a su nivel de riqueza.

Ahora bien, respecto a las variables utilizadas para el PCA, se seleccionaron 26 variables siguiendo la metodología de McKenzie (2005), las cuales pueden agruparse en tres dimensiones: infraestructura de la vivienda, acceso a servicios básicos y tenencia de activos. Todas las variables – a excepción de una (número de habitaciones en la vivienda) – son variables

dicotómicas que toman valor de 1 si el alumno tiene el activo, y 0 en caso contrario. A continuación, el detalle de las variables incorporadas:

a) Infraestructura de la vivienda:

- **Número de habitaciones:** Variable continua que indica el número de habitaciones con el que cuenta la vivienda del alumno.
- **Vivienda es propia:** Variable dicotómica que toma el valor de 1 si la vivienda del alumno es propia (con o sin título de propiedad), y 0 en caso contrario.
- **Vivienda tiene paredes de ladrillo:** Variable dicotómica que toma el valor de 1 si el material predominante en las paredes de la vivienda del alumno es de ladrillo o bloque de cemento, y 0 en caso contrario.
- **Vivienda tiene paredes de adobe:** Variable dicotómica que toma el valor de 1 si el material predominante en las paredes de la vivienda del alumno es de adobe, y 0 en caso contrario.
- **Vivienda tiene techos de concreto armado:** Variable dicotómica que toma el valor de 1 si el material predominante en los techos de la vivienda del alumno es de concreto armado, y 0 en caso contrario.
- **Vivienda tiene techos de calamina:** Variable dicotómica que toma el valor de 1 si el material predominante en los techos de la vivienda del alumno es de planchas de calamina, fibra de cemento o similares, y 0 en caso contrario.
- **Vivienda tiene pisos de cemento:** Variable dicotómica que toma el valor de 1 si el material predominante en los pisos de la vivienda del alumno es de cemento, y 0 en caso contrario.
- **Vivienda tiene pisos de tierra:** Variable dicotómica que toma el valor de 1 si el material predominante en los pisos de la vivienda del alumno es de tierra, y 0 en caso contrario.

b) Acceso a servicios básicos:

- **Vivienda tiene agua conectada a red pública:** Variable dicotómica que toma el valor de 1 si el agua que utilizan en la vivienda del alumno procede principalmente de red pública (dentro o fuera de la vivienda), y 0 en caso contrario.
- **Vivienda tiene SS.HH. conectado a red pública:** Variable dicotómica que toma el valor de 1 si el servicio higiénico de la vivienda del alumno está conectado a red pública (dentro o fuera de la vivienda), y 0 en caso contrario.
- **Vivienda tiene pozo ciego/negro:** Variable dicotómica que toma el valor de 1 si el servicio higiénico de la vivienda del alumno está conectado pozo ciego o negro, y 0 en caso contrario.
- **Vivienda tiene electricidad:** Variable dicotómica que toma el valor de 1 si la vivienda del alumno tiene alumbrado eléctrico por red pública, y 0 en caso contrario.

- **Hogar tiene teléfono fijo/móvil:** Variable dicotómica que toma el valor de 1 si en el hogar del alumno cuentan con teléfono fijo o móvil, y 0 en caso contrario.
- **Hogar tiene internet:** Variable dicotómica que toma el valor de 1 si en el hogar del alumno cuentan con internet, y 0 en caso contrario.

c) Tenencia de activos:

- **Hogar tiene equipo de sonido:** Variable dicotómica que toma el valor de 1 si en el hogar del alumno tienen un equipo de sonido, y 0 en caso contrario.
- **Hogar tiene televisor a color:** Variable dicotómica que toma el valor de 1 si en el hogar del alumno tienen televisor a color, y 0 en caso contrario.
- **Hogar tiene cocina a gas:** Variable dicotómica que toma el valor de 1 si en el hogar del alumno tienen cocina a gas, y 0 en caso contrario.
- **Hogar tiene refrigeradora:** Variable dicotómica que toma el valor de 1 si en el hogar del alumno tienen refrigeradora o congeladora, y 0 en caso contrario.
- **Hogar tiene lavadora:** Variable dicotómica que toma el valor de 1 si en el hogar del alumno tienen lavadora de ropa, y 0 en caso contrario.
- **Hogar tiene microondas:** Variable dicotómica que toma el valor de 1 si en el hogar del alumno tienen horno microondas, y 0 en caso contrario.
- **Hogar tiene licuadora:** Variable dicotómica que toma el valor de 1 si en el hogar del alumno tienen licuadora, y 0 en caso contrario.
- **Hogar tiene plancha:** Variable dicotómica que toma el valor de 1 si en el hogar del alumno tienen plancha eléctrica, y 0 en caso contrario.
- **Hogar tiene computadora:** Variable dicotómica que toma el valor de 1 si en el hogar del alumno tienen computadora, laptop o *tablet*, y 0 en caso contrario.
- **Hogar tiene automóvil:** Variable dicotómica que toma el valor de 1 si en el hogar del alumno tienen automóvil o camioneta, y 0 en caso contrario.
- **Hogar tiene motocicleta:** Variable dicotómica que toma el valor de 1 si en el hogar del alumno tienen motocicleta, y 0 en caso contrario.

En la **Tabla 13** se presentan los coeficientes del puntaje del primer componente principal del PCA realizado para cada una de las tres dimensiones por separado y utilizando todas las variables. Primero, es importante mencionar que todos los índices individuales tienen una correlación de alrededor de 0.8 con el índice de riqueza. Por otro lado, el primer componente explica el 29% del total de la varianza de las 26 variables. Además, los coeficientes son positivos para todas las variables de infraestructura, excepto para aquellas que se refieren a materiales precarios como adobe en paredes, tejas y calamina en techos, y tierra en pisos. También son positivos para todas las variables de acceso a servicios básicos, excepto para la tenencia de un sistema de saneamiento no adecuado (pozo ciego y/o negro); y de activos. En ese sentido, el primer componente del PCA parece ser un buen indicador de riqueza, al estar correlacionado

positivamente con variables que denotan un mayor bienestar y negativamente con aquellas que revelan condiciones de precariedad.

Tabla 13. Análisis de componentes principales y estadísticas descriptivas de indicadores

	Coeficientes del 1er componente				Estadísticas	
	Solo (a)	Solo (b)	Solo (c)	Todos los indicadores	Prom.	Desv. Est.
Infraestructura de la vivienda (a)						
Nro. habitaciones	0.17			0.13	2.8	1.68
Vivienda es propia	0.12			0.07	13%	0.33
Vivienda tiene paredes de ladrillo	0.47			0.26	36%	0.48
Vivienda tiene paredes de adobe	-0.32			-0.14	31%	0.46
Vivienda tiene techos de concreto armado	0.44			0.24	23%	0.42
Vivienda tiene techos de tejas	-0.16			-0.08	10%	0.31
Vivienda tiene techos de calamina	-0.25			-0.12	55%	0.50
Vivienda tiene pisos de cemento	0.38			0.21	40%	0.49
Vivienda tiene pisos de tierra	-0.45			-0.26	48%	0.50
Acceso a servicios básicos (b)						
Vivienda tiene agua conectada a red púb.		0.47		0.17	70%	0.46
Vivienda tiene SS.HH. conectado a red púb.		0.54		0.25	50%	0.50
Vivienda tiene pozo ciego/negro		-0.42		-0.16	26%	0.44
Vivienda tiene electricidad		0.37		0.16	82%	0.38
Hogar tiene teléfono fijo y/o móvil		0.30		0.16	80%	0.40
Hogar tiene internet		0.27		0.19	12%	0.33
Tenencia de activos (c)						
Hogar tiene equipo de sonido			0.30	0.21	29%	0.45
Hogar tiene televisor a color			0.33	0.24	64%	0.48
Hogar tiene cocina a gas			0.31	0.24	68%	0.47
Hogar tiene refrigeradora			0.36	0.25	32%	0.47
Hogar tiene lavadora			0.32	0.22	15%	0.36
Hogar tiene microondas			0.27	0.18	9%	0.28
Hogar tiene licuadora			0.36	0.26	44%	0.50
Hogar tiene plancha eléctrica			0.35	0.25	37%	0.48
Hogar tiene computadora			0.32	0.22	17%	0.38
Hogar tiene automóvil			0.17	0.11	5%	0.22
Hogar tiene motocicleta			0.12	0.07	13%	0.33
Eigenvalue asociado con el 1er comp.	3.25	2.48	4.40	7.53		
Prop. de la varianza asociado al 1er comp.	0.36	0.41	0.40	0.29		
Nro. variables utilizadas	9	6	11	26		
Nro. observaciones (millones)	1.8	1.8	1.8	1.8	1.8	1.8

Fuente: Base de datos armonizada. Elaboración propia.

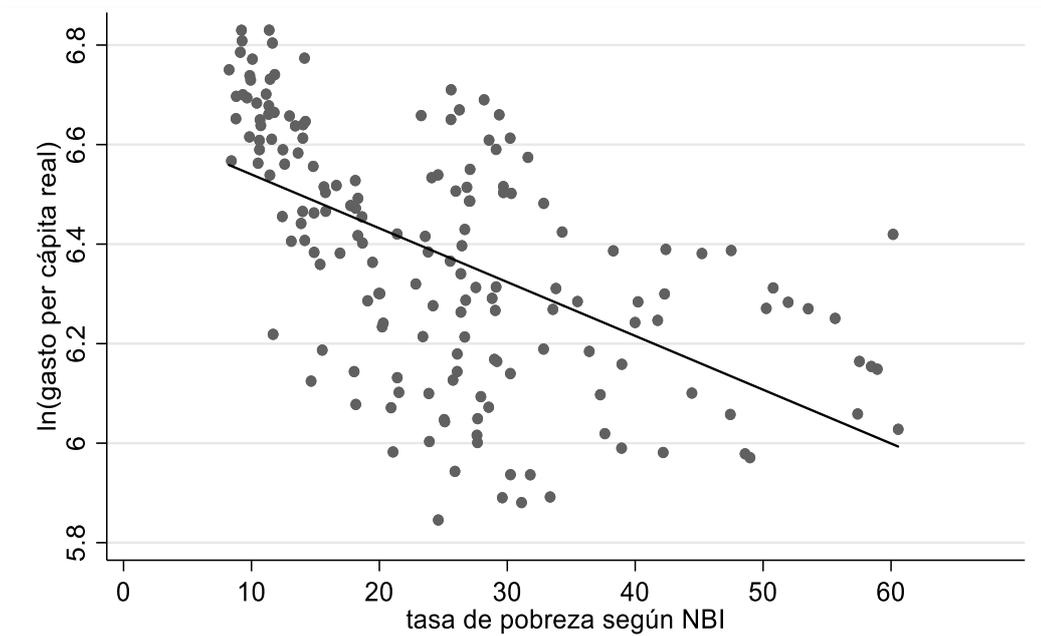
Luego de calculado el puntaje del primer componente principal a nivel de alumnos, se procedió a realizar la agregación del indicador a nivel de escuelas. Para ello, se tomó el valor promedio del puntaje del primer componente principal de los alumnos en cada escuela.

2.1.1.3. Comparación entre ambos índices

A fin de evaluar la pertinencia de los índices construidos, se replicó la metodología descrita en esta sección (método de las NBI y PCA) para todos los alumnos de las instituciones públicas del nivel de primaria identificados en la ENAHO 2008-2018. La ENAHO cuenta con todas las variables utilizadas para la construcción de ambos índices e incluye información del nivel de gasto per cápita de los hogares, lo cual permite correlacionar ambos índices construidos con esta última variable. La única dificultad es que la ENAHO no permite agregar la información a nivel de escuelas, por lo que la agregación se realizó a nivel departamental, pero manteniendo la misma metodología realizada para la construcción de los índices con data censal.

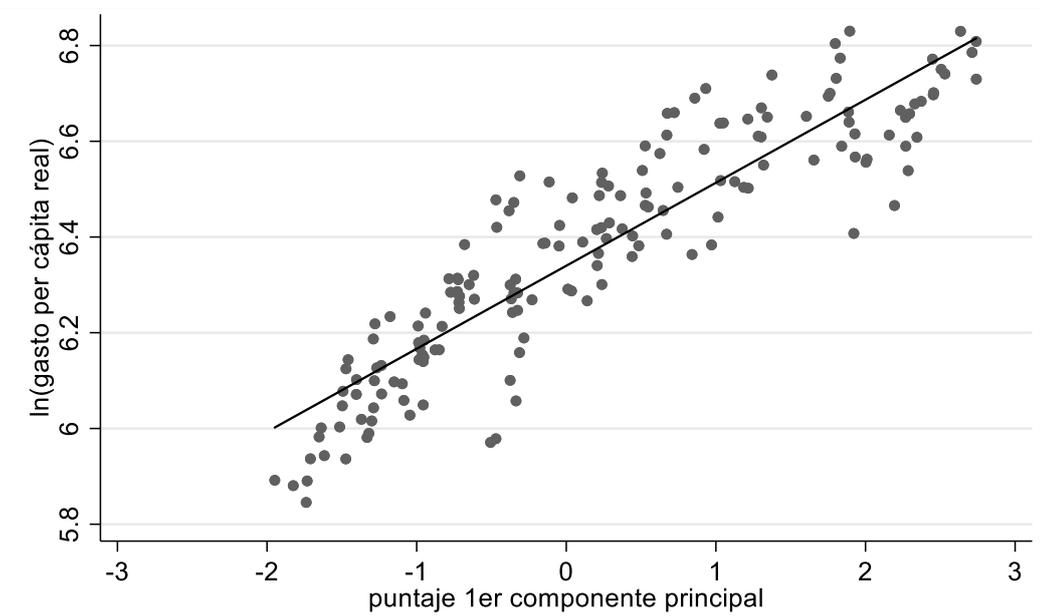
En el **Gráfico 9** y **Gráfico 10**, se presentan los resultados de la correlación entre el gasto real per cápita y el índice socioeconómico construido según el método de las NBI y el puntaje del primer componente del PCA, respectivamente. En el primer caso, se encontró una correlación de -0.56, mientras que en el segundo caso se encontró una correlación de 0.91. En términos absolutos, el puntaje obtenido a partir del primer componente del PCA se correlaciona mejor con el gasto que el índice construido a partir del método de las NBI. Esto último implica que el ordenamiento que ofrece el PCA es una mejor aproximación al gasto per cápita y, a partir de ello, a la tasa de pobreza monetaria. En consecuencia, la clasificación de escuelas que se derive de este podría tener una mejor correspondencia con el Mapa de Pobreza que elabora el INEI. Esto es particularmente importante si se considera que es altamente probable que el Mapa de Pobreza sea el principal insumo para la actualización del índice socioeconómico en el futuro.

Gráfico 9. Correlación entre el gasto real per cápita y la tasa de pobreza según NBI a nivel departamental



Fuente: ENAHO 2008-2018. Elaboración propia.

Gráfico 10. Correlación entre el gasto real per cápita y el puntaje del primer componente principal a nivel departamental



Fuente: ENAHO 2008-2018. Elaboración propia.

Ahora bien, ¿cuál de los índices socioeconómicos es mejor? Por un lado, el índice socioeconómico construido a partir de las NBI tiene la ventaja de que es sencillo de explicar, replicar y actualizar. Además, al ser una medida estandarizada de pobreza, difícilmente generará rechazos o cuestionamientos por parte de los hacedores de política pública. Sin embargo, debe señalarse que este método presenta una serie de debilidades relacionadas a la ausencia de un

sustento teórico apropiado, la arbitrariedad de sus indicadores y los umbrales críticos establecidos, su incapacidad para medir la intensidad de la pobreza, el deterioro de la representatividad de sus indicadores, entre otros (ver Feres y Mancero (2001) para una explicación más detallada). Además, otra de las desventajas de este índice es que, al medir aspectos no monetarios de la pobreza, su correlación con el gasto – aunque importante – no es tan alta.

Por su parte, el índice socioeconómico construido a partir del primer componente del PCA tiene la ventaja de que se encuentra altamente correlacionado con el gasto y, en ese sentido, no deberían existir tantas diferencias al reemplazarlo por el Mapa de Pobreza. Además, esta es una metodología muy utilizada en el país y es, por ejemplo, sobre la que descansa la construcción del Índice de Focalización de Hogares que es parte del Sistema de Focalización de Hogares (SISFOH)²⁴. Por este motivo, sugerimos utilizar el índice socioeconómico construido a partir del primer componente del PCA y es el que elegimos para ilustrar los resultados del uso conjunto del índice socioeconómico y nutricional que se presenta en la sección 5 de este informe.

2.1.2. Imputación del índice sobre escuelas faltantes

A continuación, se realizó la imputación de ambos índices socioeconómicos sobre escuelas que no contaban con información suficiente para caracterizar el estado de vulnerabilidad de las escuelas. La determinación de las necesidades de imputación se realizó a partir del balance presentado en la **Tabla 11**. Para definir que una escuela contaba con información suficiente para caracterizar el estado socioeconómico de la escuela, se consideraron dos criterios. El primero de ellos fue que cada escuela cuente con información de, por lo menos, la mitad de los alumnos de la escuela. El segundo fue que el error muestral de los indicadores construidos para cada escuela se encontrara por debajo del 10%. Para esto último, se estimó el error muestral de cada escuela, asumiendo un diseño aleatorio simple y un nivel de confianza de 10%. Luego, se considera que una escuela cuenta con información suficiente si cumple al menos uno de los dos criterios mencionados, caso contrario es objeto de imputación.

En la **Tabla 14** se presenta el detalle del ejercicio realizado. De las 29,486 escuelas de primaria que comprende el PNAEQW 2019, solo el 7% cuenta con información socioeconómica de todos los alumnos de la escuela. El 87% de las escuelas cuenta con información incompleta, pero suficiente (es decir, cumple al menos uno de los dos criterios mencionados líneas más arriba), el 4% cuenta con información incompleta e insuficiente (es decir, cuenta con información de menos de la mitad de los alumnos de la escuela y, además, presenta un error mayor a 10%), un 1% no cuenta con información socioeconómica y otro 1% corresponde a las escuelas que no lograron empatarse del padrón de PNAEQW 2019 y que, por tanto, no cuentan con información socioeconómica. Dado lo anterior, se requiere imputar el 6% del universo total (1,934 escuelas).

²⁴ Ver la Resolución Ministerial N°107-2015-MIDIS que aprueba la “Metodología para la Determinación de la Clasificación Socioeconómica”.

En breve, la imputación se realiza para tres tipos de escuelas: (i) aquellas que no contaban con información socioeconómica; (ii) aquellas que sí contaban con información socioeconómica, pero esta no era suficiente, ya sea porque contaba con información socioeconómica de menos del 50% de los alumnos de la escuela o porque el error muestral era mayor a 10%; y, (iii) aquellas que se encontraban en el directorio de PNAEQW, pero no del SIAGIE.

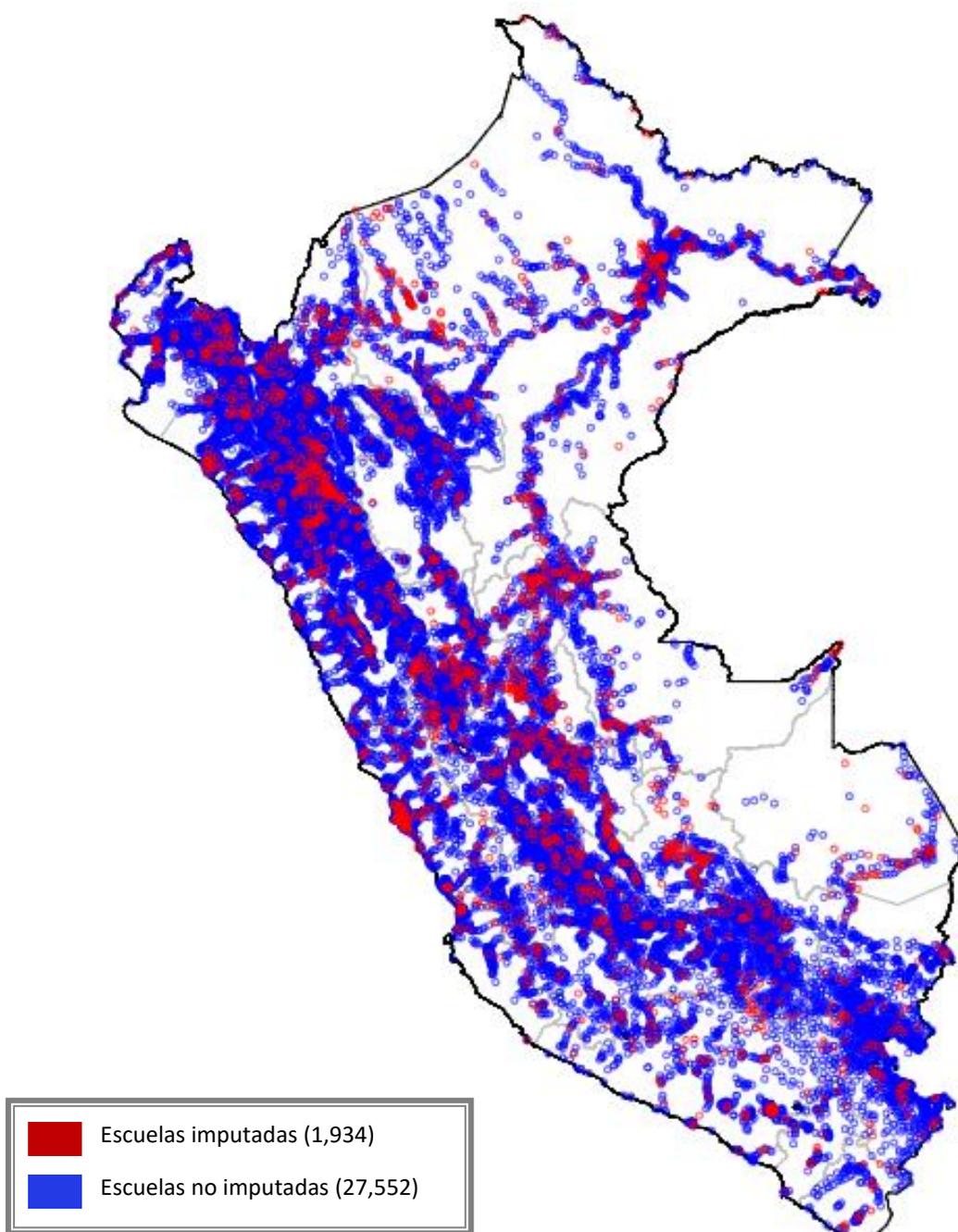
Tabla 14. Necesidades de imputación – índice socioeconómico

Situación de la información	Escuelas	
	Nro.	%
Información completa	2,006	7%
Información incompleta, pero suficiente	25,546	87%
Información incompleta e insuficiente	1,190	4%
No se tiene información	345	1%
Adicionales PNAEQW 2019	399	1%
Total	29,486	100%
Necesidad de imputación	1,934	6%

Fuente: Base de datos armonizada. **Elaboración propia.**

En el **Mapa 4** se presenta la distribución geográfica de las escuelas en el territorio nacional. En azul se encuentran las escuelas con información completa y en rojo aquellas que requirieron ser imputadas.

Mapa 4. Distribución de las escuelas según necesidades de imputación



Fuente: Base de datos armonizada. **Elaboración propia.**

La imputación se realizó a partir de la asignación de los valores promedio de las escuelas vecinas más cercanas, aprovechando la georreferenciación de las escuelas en el directorio del PNAEQW. El supuesto que subyace es que las condiciones socioeconómicas varían suavemente en el espacio, por lo que las condiciones socioeconómicas de los alumnos de las escuelas más cercanas aproximan razonablemente las condiciones socioeconómicas de una escuela sin esta información.

Operativamente se procedió del siguiente modo: para cada una de las 1,934 escuelas con información faltante se identificaron las 5 escuelas más cercanas que contaban con información suficiente y se promediaron los valores de sus índices socioeconómicos (es decir, la tasa de pobreza según NBI de la escuela y el puntaje promedio del primer componente principal obtenido en la escuela). Dicho valor promedio fue imputado a las escuelas con información faltante. La **Tabla 15** reporta la distribución de distancias promedio de las escuelas con valores imputados. Como se aprecia, las escuelas utilizadas están a 2.67 km de distancia, en promedio, y el percentil 99 se ubica a 13.16 km.

Tabla 15. Distribución de las distancias a las escuelas más cercanas

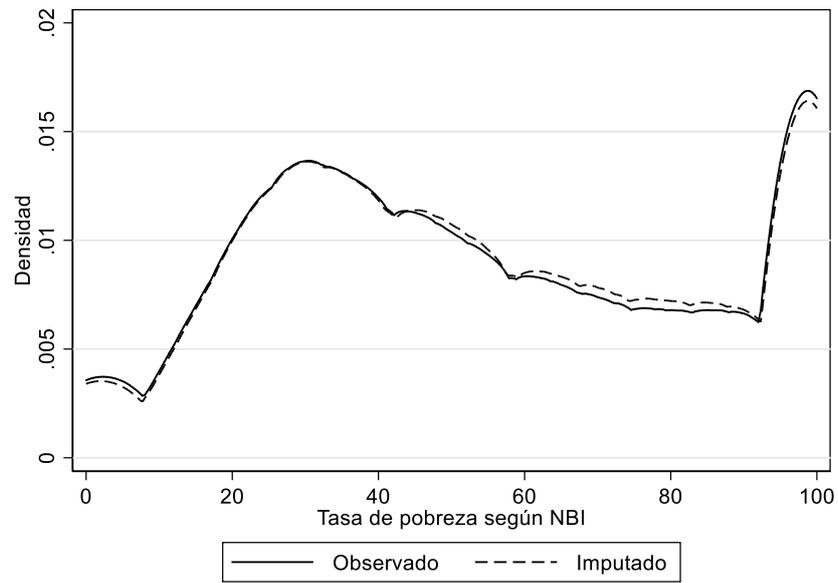
Estadístico		Km
Promedio		2.67
	1%	0.22
	5%	0.34
	10%	0.57
Percentil de distancia	25%	1.12
	50%	1.89
	75%	3.19
	90%	5.42
	95%	7.56
	99%	13.16

Fuente: Base de datos armonizada. **Elaboración propia.**

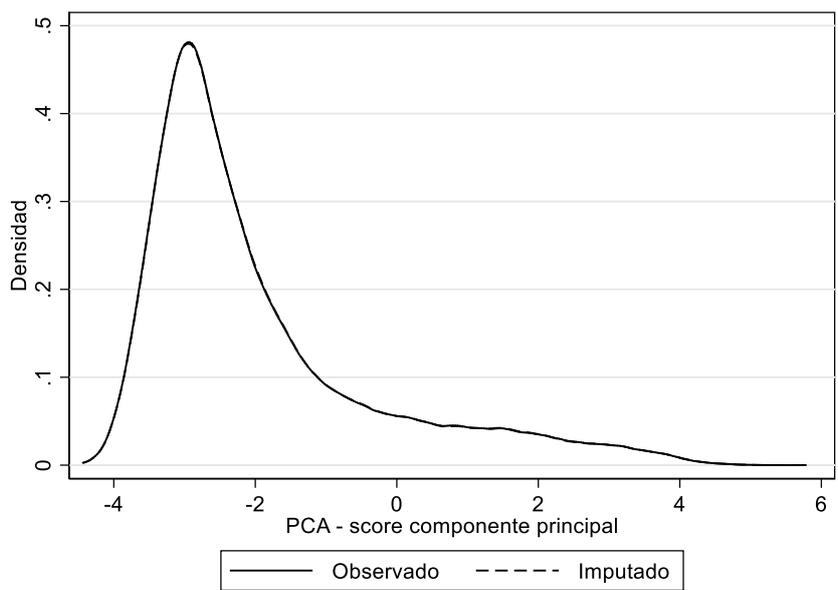
El procedimiento anterior permitió completar la información socioeconómica de todas las escuelas. En el **Gráfico 11** se presenta la distribución de los índices socioeconómicos antes de la imputación (observado) y después de la imputación (incluyendo los datos imputados). Como se observa, no se identifican diferencias importantes entre ambas distribuciones.

Gráfico 11. Distribución de los índices socioeconómicos observada e imputada

a) Índice socioeconómico según NBI



b) Índice socioeconómico según PCA



Fuente: Base de datos armonizada. **Elaboración propia.**

2.1.3. Construcción de grupos

El agrupamiento de las escuelas en grupos de necesidad alta, media y baja se realizó sobre la base de la información imputada. Se realizaron tres ejercicios de agrupamiento:

- a. **Según quintiles:** Esta es la sugerencia de la Nota Metodológica. Primero, se ponderan las escuelas según la cantidad de alumnos registrados en el padrón del PNAEQW y se dividen en quintiles según el ordenamiento que suponen los índices socioeconómicos. Este procedimiento se realiza de manera separada para ambos índices. Luego, se reagrupan los quintiles en tres grupos. El primer grupo (necesidad alta) corresponde al quintil 1, el segundo grupo (necesidad media) corresponde a los quintiles 2 y 3, y el tercer grupo (necesidad baja) a los quintiles 4 y 5.
- b. **Según quintiles + poblaciones especiales:** Es similar al ejercicio (a), pero al grupo de necesidad alta se le añaden todas las escuelas que se ubican en pueblos indígenas de la Amazonía por ser considerados de prioridad para el PNAEQW.
- c. **Análisis de conglomerados:** La clasificación de las escuelas se realiza de manera endógena a partir de un análisis por conglomerados no jerárquico por k-medias. De este modo, se agrupan las escuelas en grupos que son lo más homogéneos posibles al interior de estos, pero lo más diferentes posibles unos de otros, es decir, esta técnica permite maximizar la homogeneidad intragrupo y la heterogeneidad inter-grupo. El número de grupos se establece en tres²⁵ y los valores medios de cada uno determinan si corresponde a necesidad alta, media o baja.

En las siguientes tablas se presenta el resultado de los tres ejercicios de agrupamiento realizados para ambos índices socioeconómicos. En la **Tabla 16**, se presenta la distribución de escuelas según los distintos grupos y, en la **Tabla 17**, su equivalente en cantidad de alumnos. Como se observa, existen variaciones significativas en el agrupamiento de las escuelas bajo las metodologías empleadas. Por un lado, si bien las diferencias entre las metodologías de quintiles y quintiles más poblaciones vulnerables no son significativas, tanto a nivel de escuelas como de alumnos, sí lo son cuando se comparan con la agrupación del análisis por conglomerados. Así, por ejemplo, tomando como referencia el índice de NBI, las escuelas de alta necesidad bajo la metodología de quintiles representan el 43% del total, pero bajo la metodología de conglomerados son el 32%. En cambio, cuando se trata del índice socioeconómico, las proporciones respectivas en esta categoría son de 55% y 63%. Esta regularidad se mantiene también cuando la comparación se hace sobre alumnos. Por otro lado, a nivel de escuelas, el agrupamiento por quintiles basados en el índice NBI arroja distribuciones similares que el

²⁵ El análisis de conglomerados bajo los métodos no jerárquicos requiere generar una partición inicial de las observaciones en k conglomerados. Formalmente, este ejercicio se debe realizar para distintas particiones, de modo que luego sea posible compararlas, y se elija aquella que optimice el ejercicio. Si bien para este estudio se definieron tres grupos independientemente de si corresponden o no al número grupos óptimos, debe precisarse que esta cantidad no es aquella que maximiza el pseudo-T de acuerdo al estadístico Calinski-Harabasz, sino que la cantidad óptima de grupos estaría por sobre los cinco.

agrupamiento por conglomerados en la categoría de necesidad media. En cambio, el PCA tiende a generar un menor número de escuelas bajo el agrupamiento por conglomerados. Esta última regularidad, sin embargo, sí se mantiene a nivel de estudiantes bajo ambos índices.

Tabla 16. Clasificación de las escuelas según índices socioeconómicos

	Necesidad alta	Necesidad media	Necesidad baja	Total
Índice socioeconómico según NBI				
Quintiles	12,719	9,762	7,005	29,486
	43%	33%	24%	100%
Quintiles + poblaciones especiales	12,953	9,583	6,950	29,486
	44%	33%	24%	100%
Análisis por conglomerados	9,349	9,570	10,567	29,486
	32%	32%	36%	100%
Índice socioeconómico según PCA				
Quintiles	16,245	10,566	2,675	29,486
	55%	36%	9%	100%
Quintiles + poblaciones especiales	16,683	10,137	2,666	29,486
	57%	34%	9%	100%
Análisis por conglomerados	18,720	7,285	3,481	29,486
	63%	25%	12%	100%

Fuente: Base de datos armonizada. Elaboración propia.

Tabla 17. Clasificación de las escuelas según índices socioeconómicos – distribución de alumnos

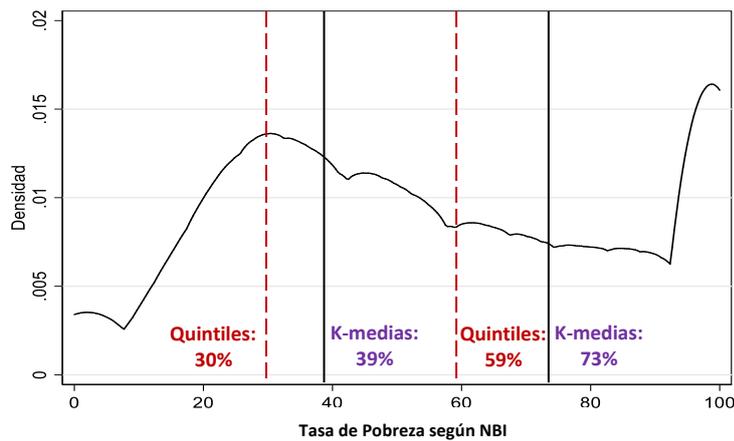
	Necesidad alta	Necesidad media	Necesidad baja	Total
Índice socioeconómico según NBI				
Quintiles	517,624	1,034,222	1,038,224	2,590,070
	20%	40%	40%	100%
Quintiles + poblaciones especiales	534,323	1,020,125	1,035,622	2,590,070
	21%	39%	40%	100%
Análisis por conglomerados	320,323	739,275	1,530,472	2,590,070
	12%	29%	59%	100%
Índice socioeconómico según PCA				
Quintiles	518,017	1,036,388	1,035,665	2,590,070
	20%	40%	40%	100%
Quintiles + poblaciones especiales	553,663	1,002,205	1,034,202	2,590,070
	21%	39%	40%	100%
Análisis por conglomerados	623,687	716,494	1,249,889	2,590,070
	24%	28%	48%	100%

Fuente: Base de datos armonizada. Elaboración propia.

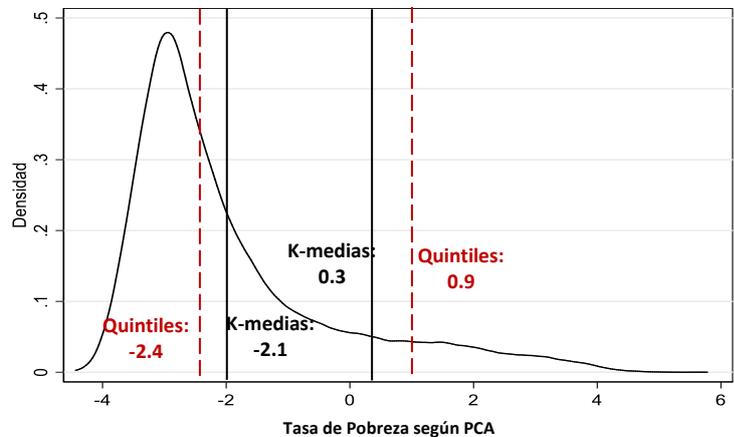
Lo anterior se explica por los umbrales estimados que definen cada grupo de necesidad bajo los dos índices (NBI y PCA) y las dos metodologías de agrupamiento (quintiles y conglomerados) (ver **Gráfico 12**). Por un lado, bajo el índice NBI, los umbrales de los conglomerados (k-medias) siempre están a la derecha de los umbrales por quintiles. En cambio, bajo el índice PCA, el umbral inferior de la categoría media del conglomerado está a la derecha del umbral inferior por quintiles, mientras que el umbral superior está a la izquierda.

Gráfico 12. Distribución de los indicadores socioeconómicos

a) Umbrales – tasa de pobreza según NBI



b) Umbrales - score primer componente PCA



Fuente: Base de datos armonizada. **Elaboración propia.**

Sobre la metodología de agrupamiento de las escuelas en la dimensión socioeconómica, el equipo consultor sugiere adoptar la estrategia basada en el análisis por conglomerados (k-medias). Esto debido a que, por un lado, reduce la discrecionalidad en el agrupamiento de escuelas dado que se basa en un procedimiento estadístico de similitud intra-conglomerado. Esta es una ventaja sobre la metodología por quintiles, que por el contrario establece cuotas fijas en cada grupo. Por otro lado, el análisis por conglomerados ofrece un agrupamiento que maximiza la homogeneidad de las escuelas al interior de cada grupo y la heterogeneidad de

estos al exterior²⁶. Los quintiles, en cambio, no necesariamente garantizan esto, debido a que las escuelas alrededor de los percentiles de corte (20 y 60) podrían ser más homogéneas entre sí que entre los grupos a los cuales se les asigna.

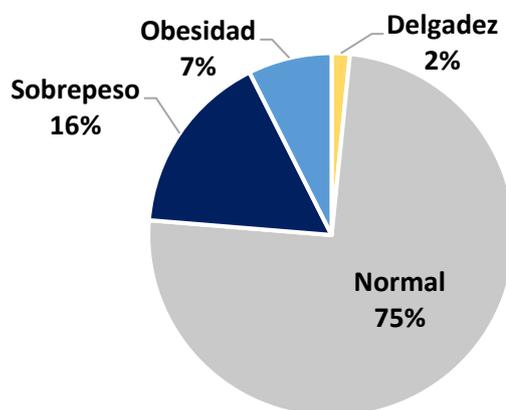
2.2. Índice nutricional

2.2.1. Construcción del índice sobre los datos disponibles

De la base de datos del PSE, se tomó la información de talla (en metros, m) y peso (en kilos, kg) y con ella se construyó el índice de masa corporal para cada estudiante (IMC), el cual se define como el peso sobre la estatura al cuadrado ($IMC = kg/m^2$). Luego, estos valores fueron comparados con los valores referenciales de la OMS para construir la variable estandarizada *z-score*^{27,28}. De 1,967,324 observaciones que tiene el PSE, se calculó el *z-score* válido para 1,755,064 estudiantes. El panel a) del siguiente gráfico presenta la distribución de este indicador. Sin embargo, solo se pudo anexar 1,049,061 estudiantes a las bases del SIAGIE.

En la medida que el interés de este índice es medir la carga nutricional de las escuelas, se inspeccionó la distribución de alumnos según su estado nutricional. Como se observa en el **Gráfico 13**, el 75% de los alumnos tiene un estado nutricional normal, el 23% tiene obesidad o sobrepeso y solo el 2% se encuentra en estado de delgadez. En la medida que la proporción de alumnos con delgadez es tan pequeña, el PMA optó por utilizar como indicador a la proporción de niños con sobrepeso u obesidad en cada escuela como índice nutricional.

Gráfico 13. Estado nutricional de los alumnos de primaria según el PSE, 2015-2016



Fuente: Base de datos armonizada. **Elaboración propia.**

Sobre las observaciones válidas, se construyó el indicador de sobrepeso y obesidad para cada estudiante (definido como una variable dicotómica que toma valor de 1 si tiene sobre

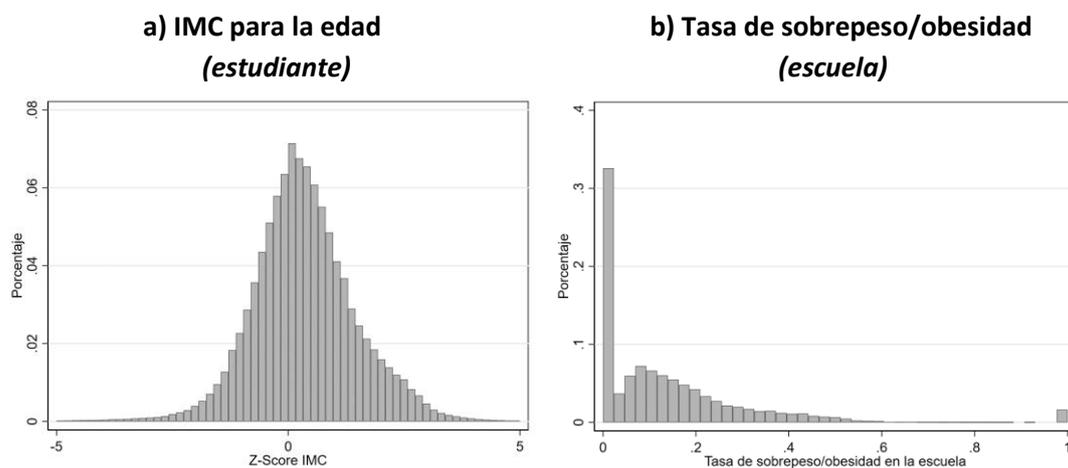
²⁶ Aún cuando en el ejercicio actual el número óptimo de conglomerados haya estado sobre 5, como se indicó en la sección 2.1.3, el procedimiento maximiza la homogeneidad dada la exigencia de establecer tres grupos.

²⁷ Ver OMS (2009), "Who AnthroPlus for Personal Computers Manual. Software for assessing growth of the world's children and adolescents". Disponible en: <https://www.who.int/growthref/tools/en/>.

²⁸ Se utilizó el paquete anthro2007 desarrollado por la OMS para el software Stata.

peso/obesidad; y 0 en caso contrario), el cual luego fue promediado a nivel de escuela. Esto permitió hallar una tasa de sobrepeso y obesidad por escuela. El panel b) del siguiente gráfico presenta esta información. No se consideró utilizar una categoría asociada a delgadez, debido a que su prevalencia en la población de estudiantes observados es reducida y ascendiente a apenas 1.7%²⁹.

Gráfico 14. Distribución de los indicadores de sobrepeso y obesidad



Fuente: Base de datos armonizada. **Elaboración propia.**

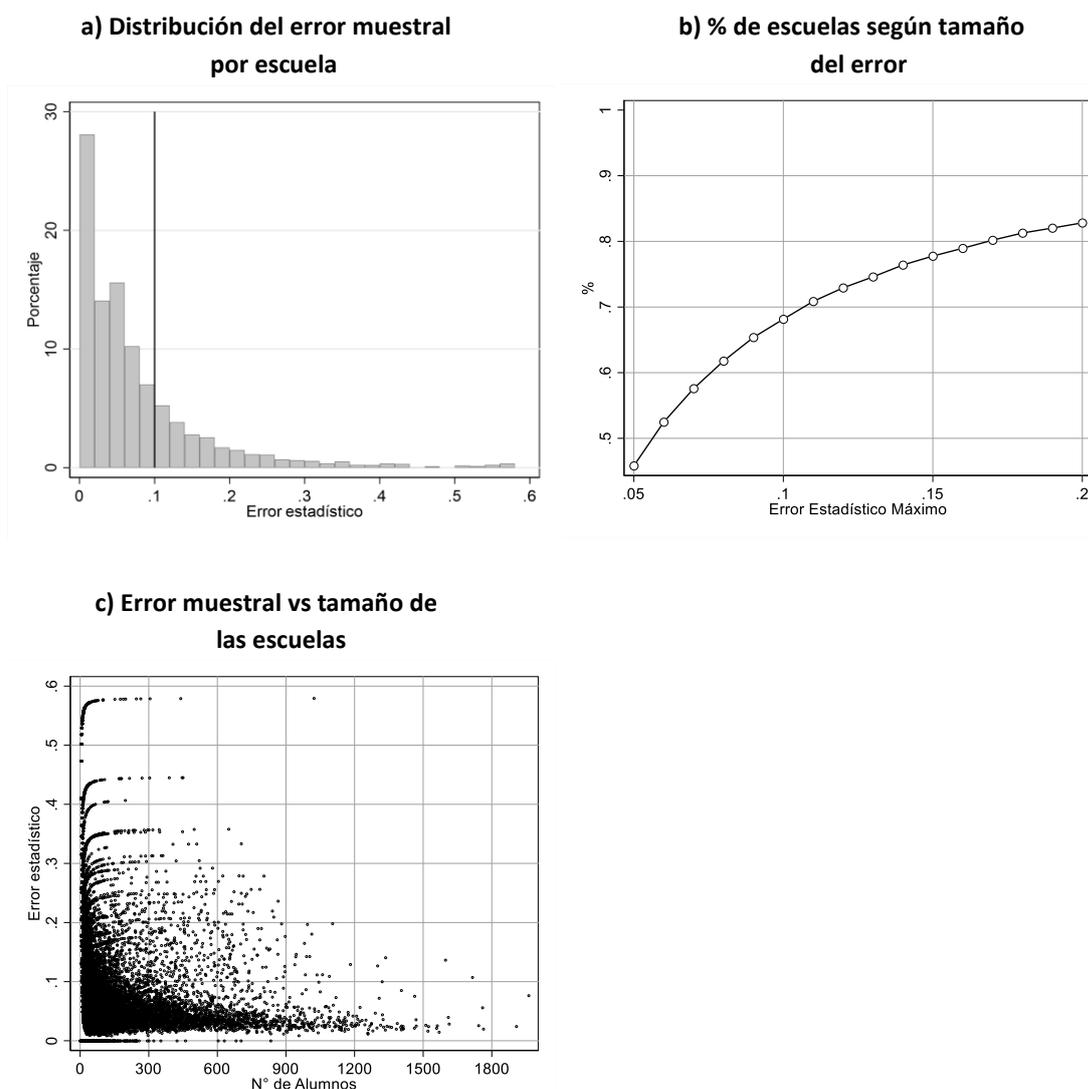
A partir de la información definida a nivel de escuela del PSE, empalmada con la información del Censo 2017, SIAGIE y el padrón del PNAEQW, se estableció una clasificación de escuelas en función a la información disponible. Para ello, inicialmente, sobre las 26,787 escuelas con información del PSE, se realizó un cálculo del error muestral por escuela asumiendo como muestra el número de alumnos con información válida del índice de masa corporal y como universo el total de estudiantes en el SIAGIE. Además, se asume un nivel de confianza de 10% y un diseño aleatorio simple. Con esto se busca identificar las escuelas con suficiente información para hacer inferencia sobre el total de la población estudiantil de primaria albergada. El supuesto que subyace es que los estudiantes con información perdida están aleatoriamente distribuidos en la escuela, por lo que utilizar la información de los encontrados permitirá una buena aproximación al resultado promedio del total de estudiantes.

En el panel a) del **Gráfico 15** se presenta la distribución del error muestral. Como se aprecia, la gran mayoría de escuelas tiene errores muestrales de pequeña magnitud, sugiriendo que en el grueso de centros educativos la información nutricional será suficiente para hacer inferencia de la tasa de sobrepeso y obesidad. El panel b) reporta el porcentaje acumulado de escuelas (con información del IMC) según el tamaño de error. Por ejemplo, poco menos de 70% tiene un error menor de 10% y poco menos de 80%, menor a 15%. El panel c) del mismo gráfico presenta el error muestral según el tamaño del centro educativo, medido como número de alumnos. Claramente, se aprecia que las escuelas más pequeñas son las que tienen los errores muestrales más altos.

²⁹ Las tasas estimadas en la base fueron las siguiente: delgadez 1.6%, estado normal 74.6%, sobrepeso 16.3% y obesidad 7.4%.

En lo que sigue se entenderá como escuelas con información muestral suficiente en la variable nutricional aquellas que tienen un error inferior a 10%. Para estas, se entenderá que la información es suficiente para hacer estimaciones confiables de la prevalencia de sobrepeso y obesidad en la escuela.

Gráfico 15. Distribución de los indicadores de sobrepeso y obesidad



Fuente: Base de datos armonizada. **Elaboración propia.**

Realizados los cálculos anteriores, se procedió a clasificar a las escuelas en función de la información disponible, para lo cual cada centro educativo se vinculó con los resultados del análisis muestral de la sección del índice socioeconómico. El panel superior de la **Tabla 18** presenta el detalle de la información. Allí, las filas hacen referencia a las escuelas registradas en el padrón del PNAEQW que tienen información completa (es decir, todos los alumnos del SIAGIE cuentan con medición nutricional), las que tienen información simultánea del PSE, Censo y SIAGIE y un error estimado asociado a la tasa de alumnos con sobrepeso y obesidad menor a 10% y mayor o igual a 10%; las que solo tienen información del SIAGIE y el Censo 2017, y las que

únicamente están registradas en el padrón del PNAEQW pero no figuran en SIAGIE. En las columnas de la tabla se presenta aquellas que bajo el análisis muestral del índice socioeconómico tienen un error estimado menor y mayor o igual a 10% (y, por tanto, con la misma consideración que con el índice nutricional, no requieren imputación).

A partir del panel superior, las escuelas han sido clasificadas en función de la suficiencia de información para efectos del cálculo del índice nutricional (panel inferior).

- **Escuelas con información completa del PSE:** Son aquellas en las que todos los alumnos del SIAGIE tienen información nutricional en el PSE.
- **Escuelas con información incompleta pero suficiente del PSE:** Son las que tienen información completa del PSE y un error estimado inferior a 0.1.
- **Escuelas con información incompleta e insuficiente del PSE:** Son aquellas que, por un lado, tienen información insuficiente del PSE (es decir, un error igual o superior a 0.1) y las que carecen de información del PSE, pero tienen información suficiente del censo (es decir, un error inferior a 10% bajo la clasificación de la sección socioeconómica) y, por otro, las que carecen de información del PSE y tiene información también insuficiente del censo (error igual o superior a 10%).
- **Escuelas adicionales de PNAEQW:** Son las escuelas que figuran en el padrón del PNAEQW, pero no están registradas en el padrón del SIAGIE de 2017.

Tabla 18. Clasificación de la información a nivel de escuelas

Grupo	Análisis socioeconómico		Total
	e<0.1	e>=0.1	
Detalle de las escuelas			
PNAEQW + Información completa	617 [A]	-	617
PNAEQW + SIAGIE + Censo 2017 + PSE(e<0.1)	18,900 [B]	579 [E]	19,479
PNAEQW + SIAGIE + Censo 2017 + PSE(e>=0.1)	6,235 [C]	456 [F]	6,691
PNAEQW + SIAGIE + Censo Nacional L 2017	1,800 [D]	155 [G]	1,955
PNAEQW + SIAGIE	-	345 [H]	345
Solo PNAEQW	-	399 [I]	399
Total	27,552	1,934	29,486
Clasificación según suficiencia de la información			
i. Con información completa del PSE	[A]		617
ii. Con información incompleta pero suficiente del PSE	[B]+[E]		19,479
iii. Con información incompleta e insuficiente del PSE	[C]+[D]+[F]+[G]+[H]+[I]		8,991
iv. Con datos insuficiente del PSE, pero suficientes del Censo 2017	[C]+[D]		8,035
v. Sin datos o con datos insuficiente del Censo 2017	[F]+[G]+[H]		956
vi. Adicionales PNAEQW	[I]		399
Total			29,486

Fuente: Base de datos armonizada. **Elaboración propia.**

Tomando en cuenta la clasificación de la información del panel inferior de la **Tabla 18**, se ha planteado la siguiente estrategia de construcción del índice nutricional por escuela (porcentaje de estudiantes con sobrepeso y obesidad) que considera tres etapas (ver **Tabla 19**):

- **Primera etapa:** Para aquellas con información completa del PSE y para aquellas con información incompleta pero suficiente, se hará un cálculo directo. Es decir, se tomará directamente el porcentaje de estudiantes con sobrepeso y obesidad.
- **Segunda etapa:** Para aquellas con información insuficiente del PSE, pero suficiente del CENSO 2017, se hará un proceso de imputación a partir de un modelo de pronóstico basado técnicas de *machine learning*.
- **Tercera etapa:** Para aquellas con sin datos del PSE y datos insuficientes del Censo 2017, más las escuelas adicionales de PNAEQW, se implementará un proceso de imputación basado en la cercanía geográfica de las escuelas con información completa o reconstruida en la etapa anterior.

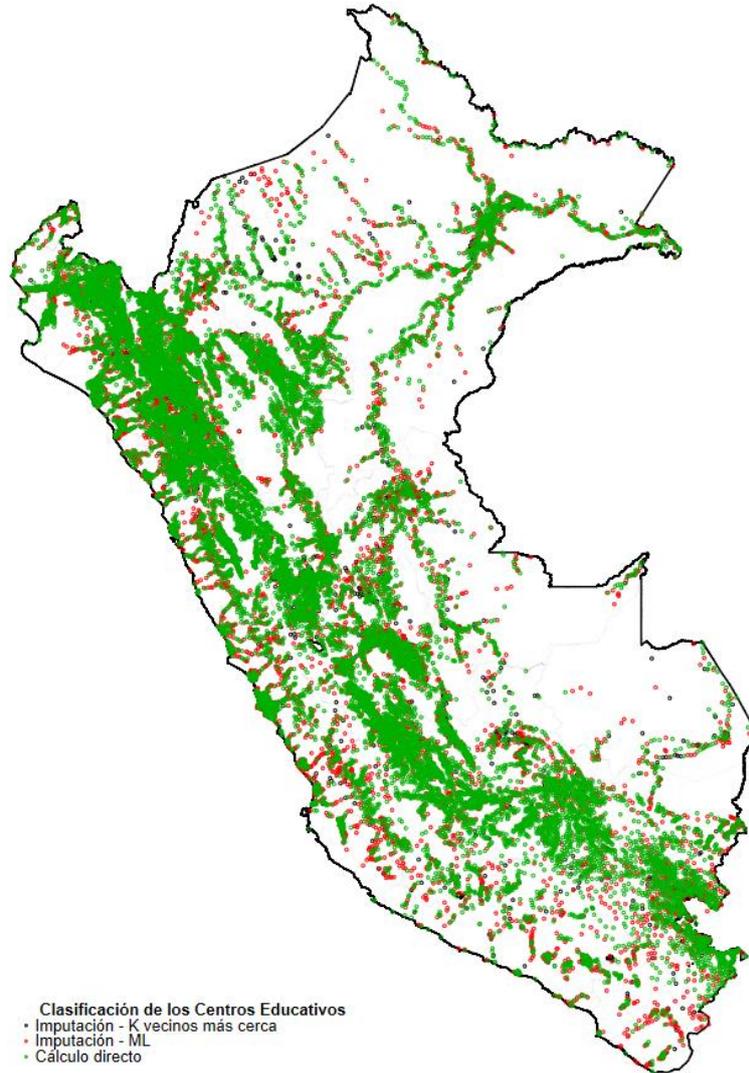
Tabla 19. Clasificación de las escuelas según métodos de cálculo del indicador nutricional

Muestra		Grupos	#	Método
	Total de escuelas		29,486	.-.
A	Con información completa y/o suficiente del PSE	(i)+(ii)	20,096	Cálculo directo
B	Con información insuficiente del PSE, pero suficiente del Censo 2017	(iv)	8,035	Imputación a través de <i>machine learning</i>
C	Sin datos o con datos insuficientes del Censo 2017 más adicionales de PNAEQW	(v)+(vi)	1,355	Imputación a través del vecino más cercano

Fuente: Base de datos armonizada. **Elaboración propia.**

El siguiente mapa presenta la distribución de escuelas según los tres grupos de clasificación. Claramente se aprecia que las escuelas con información por imputar están distribuidas a lo largo de todo el espacio nacional.

Gráfico 16. Distribución de las escuelas según método de cálculo



Fuente: Base de datos armonizada. **Elaboración propia.**

2.2.2. Imputación del índice sobre escuelas faltantes

2.2.2.1. Imputación a través de *machine learning*

A través de este procedimiento apuntamos a recuperar el indicador nutricional sobre las escuelas que cuentan con información insuficiente del PSE, pero tienen información suficiente de diversas variables socioeconómicas del censo. Para ello, se ha estimado un modelo de pronóstico basado en técnicas de *machine learning* (Richman et al 2008).

Para efectos de la estimación³⁰, se utilizó la técnica del LASSO (*least absolute shrinkage and selection operator*, por sus siglas en inglés), el cual se basa en la resolución del siguiente problema de optimización con respecto a β y para un parámetro $\lambda > 0$:

$$\beta_{lasso}(\lambda) = \arg \min \frac{1}{n} \sum_{i=1}^n (y_i - x_i' \beta)^2 + \frac{\lambda}{n} \sum_{j=1}^p \phi_j |\beta_j| \quad (i)$$

Donde y_i es la variable dependiente a nivel de escuela i (la proporción de alumnos con sobrepeso y obesidad), x_i es un vector de variables explicativas y β y ϕ_j son parámetros por estimar. En esta función, el primer sumando penaliza la falta de ajuste, tal como lo hace mínimos cuadrados ordinarios. El segundo sumando, llamada función de regularización, en cambio, penaliza la inclusión de variables irrelevantes. De esta manera, el estimador busca maximizar el ajuste del modelo (con el primer sumando), pero minimizando el número de regresores incluidos (segundo sumando).

La función de regularización depende de dos parámetros. El primero, λ , controla la penalización global de los parámetros incluido en β , mientras que ϕ_j penaliza la inclusión de cada parámetro en específico.

Un punto clave en estas técnicas es la selección del parámetro de penalización global λ . En la práctica, se propone una serie de valores en orden decreciente y se elige aquél que arroje un modelo con mejor ajuste a los datos. Para esto, existen diversas técnicas:

- **Técnicas basadas en criterios de información:** Los criterios de información son indicadores basados en propiedades de pérdida de eficiencia y consistencia en la selección del modelo. Existen diversos criterios de información:
 - **Akaike:** $AIC(\lambda) = n \log(\sigma^2(\lambda)) + 2gl(\lambda)$
 - **Bayesiano:** $BIC(\lambda) = n \log(\sigma^2(\lambda)) + 2gl(\lambda) \log(n)$
 - **Akaike corregida por sesgo:** $AIC_c = n \log(\sigma^2(\lambda)) + 2gl(\lambda) \frac{n}{n-gl(\lambda)}$
 - **Bayesiano Extendido:** $EBIC(\lambda) = n \log(\sigma^2(\lambda)) + 2\xi gl(\lambda) \log(p)$

Donde n es el número de observaciones, σ^2 es el residuo del error estimado en un valor dado de λ , gl son los grados de libertad, y $\xi \in [0,1]$ es un parámetro arbitrario.

- **Técnicas basadas en validación cruzada:** La idea detrás de esta técnica es la de estimar el modelo en una porción de la muestra y validar el ajuste generado por los parámetros en la porción de muestra restante. Es decir, evaluar el ajuste fuera de la muestra tomando como indicador el error cuadrático medio de predicción (ECMP).

³⁰ La exposición de la técnica de LASSO se basa en Hastie, T.; R. Tibshirani y J. Friedma (2008); C. Hansen y M. Schaffer (2019), entre otros.

Para implementar esta técnica, inicialmente la muestra debe partirse en \mathcal{K} porciones de igual tamaño. Luego, denotando por \mathcal{K}_k al conjunto de observaciones correspondiente a la k -ésima submuestra y n_k al número de observaciones correspondiente, para un valor dado de λ , las estimaciones deben hacerse en las $\mathcal{K} - 1$ submuestras (que se llamarán muestras de entrenamiento) y la validación en la submuestra \mathcal{K} restante. Luego, el error cuadrático medio será construido del siguiente modo:

$$ECMP_k(\lambda) = \frac{1}{n_k} \sum_{i \in \mathcal{K}_k} (y_i - x_i' \beta_k(\lambda))^2$$

Se elegirá el parámetro λ que minimice el error cuadrático medio después de realizada todas las iteraciones. Es recomendable que luego de cada iteración, la selección de la porción \mathcal{K}_k utilizada para validación rote entre todas las submuestras.

Con las consideraciones anteriores, las estimaciones para la elección del mejor modelo predictivo se hicieron sobre la muestra A de la **Tabla 19**. Esta se dividió aleatoriamente en dos grupos de igual tamaño. En el primer grupo se estimó el modelo (i) y en el segundo se estimaron los errores cuadráticos medios de pronóstico fuera de la muestra. Estos últimos no deben confundirse con los ECMP del estimador de Validación Cruzada del Lasso, que son más bien parte del procedimiento para la selección del λ óptimo. El ejercicio de cálculo del ECMP al que aquí nos referimos ha sido implementado después de realizar las estimaciones con todas las técnicas descritas líneas arriba, y apuntan a identificar la técnica que arroja el mejor modelo de pronóstico. A modo de comparación, también se ha implementado una estimación por Mínimos Cuadrados Ordinarios.

Como se indicó anteriormente, en todas las estimaciones la unidad de observación ha sido la escuela, teniendo como variable dependiente la proporción de estudiantes con sobrepeso y obesidad. Las variables independientes han sido el número de usuarios del PNAEQW, su cuadrado, la proporción de estudiantes varones en la escuela, la proporción de alumnos en cada grado de primaria, la proporción de niños que trabajan, características de los miembros del hogar de los estudiantes (el número promedio de miembros, el promedio de miembros de 6 a 12 años de edad, de 12 a 17 años de edad, la proporción de estos que asisten a la escuela, el número de miembros que pertenecen a la población en edad de trabajar (PET), el número promedio de ocupados, la proporción de jefes de hogar en cada nivel educativo, entre otros), características de la vivienda (paredes de ladrillo o piedra con cemento, de adobe o tapia, u otros más precarios; techos de cemento, tejas, madera, calamina y otro; pisos con material diferente a tierra), acceso a servicio (agua, desagüe, electricidad, combustible para cocinar), propiedad de la vivienda, acceso a activos (electrodomésticos, vehículos, entre otros), la altitud de la escuela (en metros sobre el nivel del mar), la distancia hacia vías nacionales, departamentales y vecinales más cercanas, y la región natural de ubicación de los centros educativos (costa, sierra y selva).

Los resultados se presentan en la **Tabla 20**. A modo de ejemplo, el mejor modelo seleccionado bajo el criterio *AIC* (columna 1) contiene 101 variables explicativas, reporta un ECMP fuera de

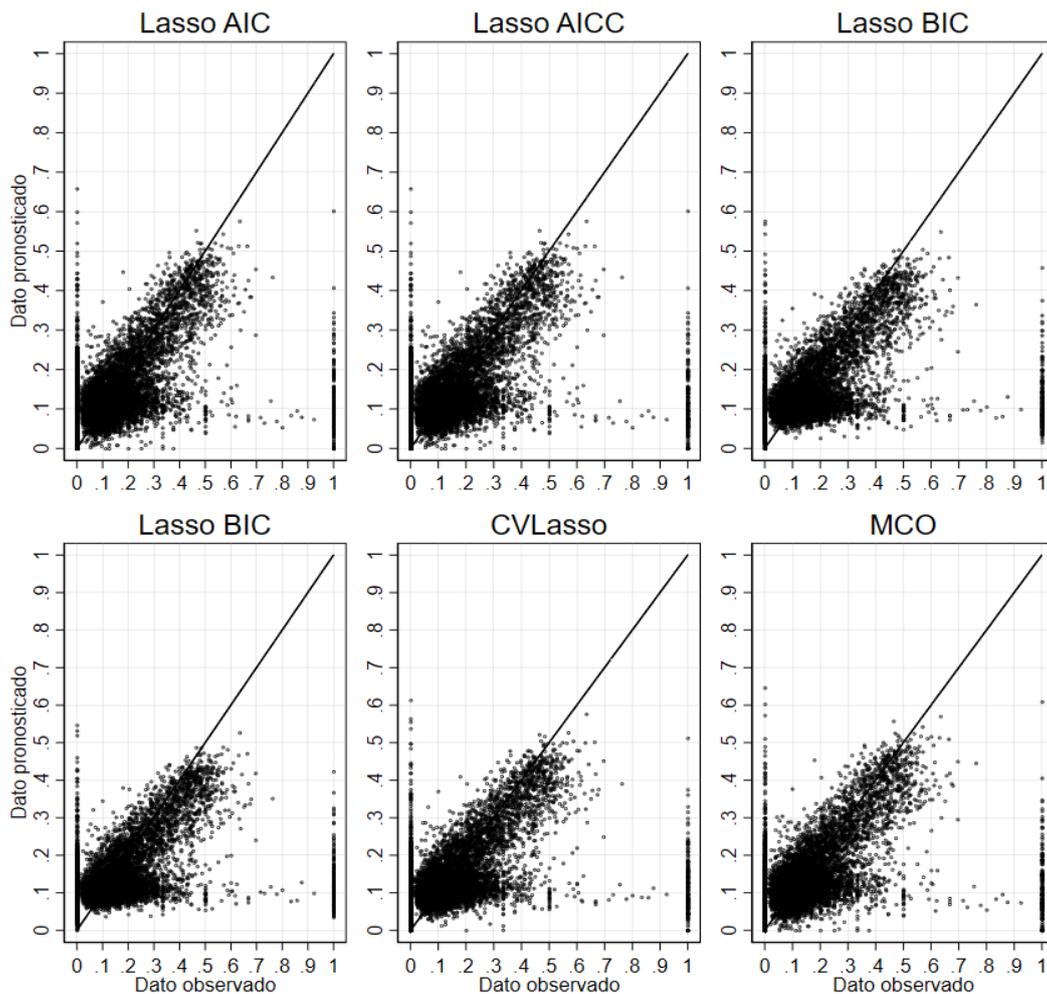
la muestra de 0.153 y un ECMP dentro de la muestra de 0.142. Estos resultados son similares bajo el criterio AIC_c . En cambio, bajo el criterio $EBIC$ el modelo óptimo tiene 23 variables y un ECMP fuera de la muestra de 0.152. En el **Gráfico 17** se presentan los pronósticos fuera de la muestra de cada estimador (es decir, sobre el grupo de escuelas no utilizadas en la construcción de los modelos) y se los compara con los datos observados. El ajuste, como se aprecia, es similar en todos los casos. De hecho, la correlación cruzada entre las variables pronosticadas fuera de la muestra es superior a 96% en el menor de los casos. A partir la **Tabla 20**, se elige como mejor modelo de pronóstico el modelo Lasso con validación cruzada (columna 5). Este tiene el menor ECMP fuera de la muestra (0.152), junto con los modelos BIC y EBIC (columnas 3 y 4, respectivamente), y entre estos tiene el menor ECMP dentro de la muestra.

Tabla 20. Resultados de los modelos de estimación

	Lasso					MCO
	AIC	AIC_c	BIC	EBIC	VC	
	[1]	[2]	[3]	[4]	[5]	
# escuelas	10,013	10,013	10,013	10,013	10,013	10,013
ECMP fuera de la muestra	0.153	0.153	0.152	0.152	0.152	0.153
ECMP dentro de la muestra	0.142	0.142	0.144	0.145	0.143	0.142
# Variables	101	101	31	23	56	124

Fuente: Base de datos armonizada. **Elaboración propia.**

Gráfico 17. Pronóstico fuera de la muestra bajo cada modelo



Fuente: Base de datos armonizada. **Elaboración propia.**

Con el mejor modelo de proyección seleccionado anteriormente, se hizo el pronóstico de la tasa de sobre peso y obesidad para cada escuela de la muestra B de la **Tabla 19** (es decir, para las que carecen de información del PSE o tiene información insuficiente, pero cuenta con regresores suficientes del censo de población y vivienda).

2.2.2.2. Imputación a través del vecino más cercano

La imputación del vecino más cercano es similar a la estrategia de imputación del índice nutricional. Es decir, tomando ventaja de las coordenadas de ubicación georeferencial de las escuelas provista por el PNAEQW, para cada escuela del grupo C de la **Tabla 19**, se identificó a hasta cinco escuelas más cercanas con información disponible (obtenida a partir de cálculo directo o a partir de la imputación descrita anteriormente). La distribución de distancias de las escuelas más cercanas se presenta en la **Tabla 21**. Como se aprecia, el promedio es de 2.44km y el percentil 99 de solo 11.8km.

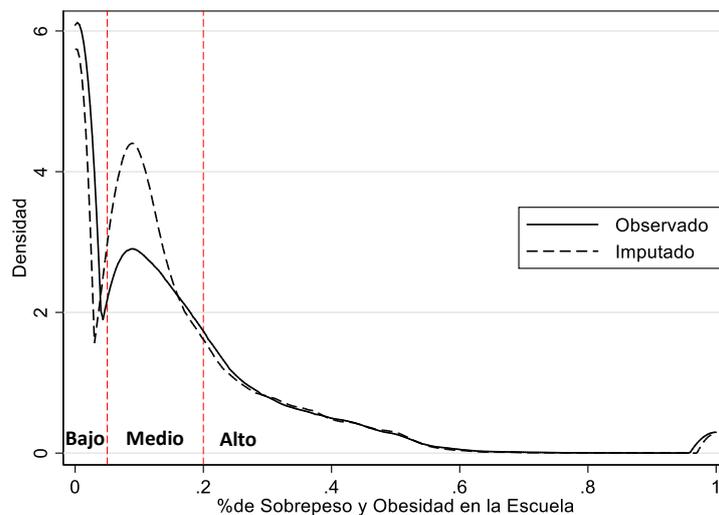
Tabla 21. Distribución de las distancias a las escuelas más cercanas

Estadístico		Km
Promedio		2.44
	1%	0.19
	5%	0.29
	10%	0.41
Percentil de distancia	25%	0.95
	50%	1.71
	75%	2.98
	90%	5.15
	95%	7.34
	99%	11.80

Fuente: Base de datos armonizada. **Elaboración propia.**

Para cada escuela del grupo C, se promediaron los porcentajes del índice nutricional de las escuelas más cercanas. Este, finalmente, fue el valor imputado sobre las 1,355 escuelas con información faltante. El **Gráfico 18** presenta la distribución empírica de la variable nutricional observada e imputada.

Gráfico 18. Distribución de la tasa de alumnos con sobrepeso y obesidad observada e imputada



Fuente: Base de datos armonizada. **Elaboración propia.**

Finalmente, el índice nutricional fue validado a partir de información primaria. En el **Anexo 12.2** se presenta el detalle del procedimiento realizado.

2.2.3. Construcción de grupos

Hasta este punto, ya se tiene toda la información construida del indicador de sobrepeso y obesidad en la escuela. El siguiente paso es agrupar a las escuelas. Para ello, la nota metodológica establecía una clasificación en base a dos parámetros: alta/baja prevalencia de desnutrición y alta/baja prevalencia de sobrepeso y obesidad. Sin embargo, el indicador de desnutrición no es factible de ser construido para el segmento etario de estudiantes considerados en el estudio (primaria), por lo que solo se debió clasificar a las escuelas en base al único índice nutricional. Por recomendación del PMA, se establecieron tres categorías relevantes basados en convenciones internacionales de salud pública en presencia de tasas altas de anemia³¹. Las categorías son:

- Prevalencia \geq 20% (alto sobrepeso y obesidad en la escuela)
- 5% \geq Prevalencia < 20% (sobrepeso y obesidad media en la escuela)
- Prevalencia < 5% (bajo sobrepeso y obesidad en la escuela)

En la **Tabla 22** se presentan los resultados. Considerando la totalidad de los datos, el 28% de las escuelas tendría baja prevalencia de sobrepeso y obesidad, el 28% tendría una prevalencia media y el 23.5%, una prevalencia alta. Cuando se observa a nivel de estudiantes (aproximado por el número de usuarios de PNAEQW el año 2019), sin embargo, las proporciones se concentran en población con alta prevalencia de sobrepeso y obesidad. Así, de los casi 2.6 millones de alumnos considerados, 1.5 millones (cerca de 60%) se encuentra en esta categoría, y 824 mil (alrededor de 32%) en la categoría de prevalencia media. solo 224 mil (poco menos de 9%) se encuentra en la categoría de baja prevalencia.

Tabla 22. Clasificación de las escuelas según el índice nutricional

Nivel de agregación	Prevalencia de sobrepeso/obesidad			Total
	Alta (>20%)	Media (5% a <20%)	Baja (<5%)	
Escuelas	6,974	14,266	8,246	29,486
	24%	48%	28%	100%
Usuarios	1,548,296	828,238	213,536	2,590,070
	60%	32%	8%	100%

Fuente: Base de datos armonizada. Elaboración propia.

³¹ Según el Plan Nacional para la reducción de la Anemia 2017-2021 elaborada por el Ministerio de Salud, la Organización Mundial de la Salud clasifica la anemia como un problema de salud pública de la siguiente manera: 5%, no representa un problema de salud pública; 5 a 19.9%, problemas de salud pública leve; 20 a 39%, problema de salud pública moderado; 40%, problema de salud pública grave.

12.2. VALIDACIÓN DE LOS ÍNDICES DE PRIORIZACIÓN

De acuerdo con lo señalado en la Nota Metodológica, la metodología empleada para la construcción del índice de priorización debe ser validada con información primaria. Para ello, se sugiere, primero, recoger información primaria de una muestra de 384 estudiantes de primero de primaria en cada región geográfica del país (costa, sierra y selva); luego, replicar la metodología empleada para la construcción del índice de priorización utilizando información primaria; y, finalmente, comparar ambos índices y determinar la validez de la metodología implementada. De acuerdo con los TdR de este estudio, esta comparación consiste en un análisis de la correlación que existe entre la información construida a partir de información secundaria con la real. En respuesta a lo solicitado, el equipo consultor validó la metodología empleada pero solo para la construcción del índice nutricional. Esto porque toda la información utilizada para la construcción del índice socioeconómico corresponde a información censal³² y, en ese sentido, no requiere de validación alguna.

El levantamiento de información antropométrica se llevó a cabo entre octubre y noviembre de 2018. Durante este periodo, se recogió la talla y el peso de 1,209 niños de primero de primaria en tres regiones del país³³ (412 en Lima Metropolitana, 381 en Huancavelica y 416 en Loreto) en 192 instituciones educativas públicas (64 escuelas por departamento). El detalle del operativo de campo se puede encontrar en el segundo entregable de este estudio (“Informe de campo”). Luego, se replicó la metodología utilizada para construir el índice nutricional con la información del PSE. En breve, esta consiste en calcular el IMC de los niños y computar sus z-scores a partir de los patrones de referencia de la OMS para niños de 5 a 19 años. Luego, dichos z-scores se comparan con los siguientes puntos de corte para identificar el estado nutricional del niño: delgadez, normal, sobrepeso u obesidad. Finalmente, se estima la proporción de alumnos con sobrepeso u obesidad en la escuela, dando origen al índice nutricional.

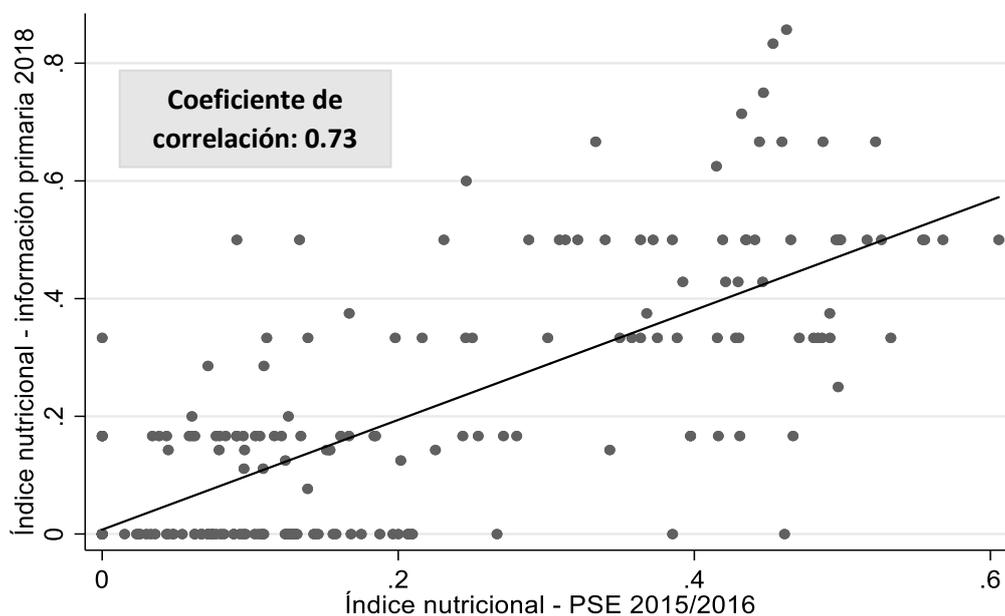
A continuación, analizamos la correlación entre el índice nutricional construido a partir de la información recogida en 2018 y el índice construido a partir de la información del PSE de los años 2015 y 2016. Como se observa en el **Gráfico 19**, la correlación entre ambos índices es positiva y asciende a 0.73. Asimismo, estimamos dos modelos para analizar el grado de ajuste de ambos índices. El primero regresiona el índice construido con información primaria con el construido con el PSE; y el segundo incorpora variables dicotómicas por departamento a lo anterior. Como se observa en la **Tabla 23**, en ambos casos, el coeficiente del índice nutricional es positivo (de 0.933 en primer modelo y de 0.547 en el segundo) y significativo, lo que implica que el índice construido se encuentra altamente correlacionado con la información real. De este

³² En la Nota Metodológica se sugiere validar el índice socioeconómico. No obstante, esta sugerencia corresponde a un contexto en el que se utiliza la información muestral de la ENAHO (gasto per cápita de los hogares de los alumnos de las escuelas) para predecir la condición socioeconómica de las escuelas. Al descartar el uso de esta variable y, en su lugar, utilizar indicadores medidos a partir del Censo Nacional, la validación de este índice se vuelve innecesaria.

³³ La selección de departamentos proviene de la Nota Metodológica: Lima Metropolitana (costa), Huancavelica (sierra) e Iquitos (selva).

modo, no solo validamos la metodología utilizada para la construcción del índice nutricional, sino que la información del 2015/2016 puede utilizarse para caracterizar el estado nutricional de los niños tres años más tarde.

Gráfico 19. Correlación entre índices nutricionales



Fuente: Base de datos armonizada. **Elaboración propia.**

Tabla 23. Modelos de ajuste de los índices nutricionales

	Modelo 1	Modelo 2
Índice nutricional – PSE	0.933***	0.547***
Lima Metropolitana	No	Sí
Huancavelica	No	Sí
Loreto	No	Sí
Constante	0.007	0.044
Obs.	192	192
R cuadrado	0.545	0.574

Fuente: Base de datos armonizada. **Elaboración propia.**

Nota: En el Modelo 1, la variable dependiente es el índice nutricional construido con información primaria y la variable independiente es el índice nutricional construido con información del PSE. En el Modelo 2 es similar al anterior, pero incorpora variables dicotómicas departamentales.

12.3. ARCHIVOS DIGITALES

Adjunto al presente informe, con el siguiente contenido:

- 001 Entregable 4: Informe final
- 002 Base de datos del estudio + dofiles
- 003 Presentación ejecutiva en formato en Power Point