



UNIVERSIDAD
DE PIURA

FACULTAD DE CIENCIAS ECONÓMICAS Y EMPRESARIALES

**El impacto de la malnutrición de los estudiantes en su
rendimiento escolar: Un estudio en niños y adolescentes
peruanos**

Tesis para optar el Título de
Economista

**Mauricio Andre Beuzeville Montañez
Macarena Ayleen Mansilla Mahmud**

**Asesora:
Ph.D. Celia Patricia Vera Rojas**

Lima, junio de 2025

Declaración Jurada de Originalidad del Trabajo Final

Yo, Mauricio Andre Beuzeville Montañez, egresado del Programa Académico de Economía de la Facultad de Ciencias Económicas y Empresariales de la Universidad de Piura, identificado con DNI: 71888762, declaro que:

Soy autor del trabajo final titulado:

“El impacto de la malnutrición de los estudiantes en su rendimiento escolar: Un estudio en niños y adolescentes peruanos”

El mismo que presento bajo la modalidad de Tesis para optar el Título profesional de Economista.

Que el trabajo se realizó en coautoría con los siguientes alumnos de la Universidad de Piura.

- Macarena Ayleen Mansilla Mahmud, identificado con DNI: 72970452

El texto de mi trabajo final es original y no vulnera los derechos de terceros o, de ser el caso, derechos de los coautores, incluidos los derechos de propiedad intelectual, datos personales, entre otros. En tal sentido, el texto de mi trabajo final no ha sido plagiado total ni parcialmente, para lo cual, he respetado las normas internacionales de citas y referencias de las fuentes consultadas. Asimismo, el texto del trabajo final que presento no ha sido publicado ni presentado antes en cualquier medio electrónico o físico; y que la investigación, los resultados, datos, conclusiones y demás información presentada que atribuyo a mi autoría son veraces.

En caso de detectarse el incumplimiento de lo declarado asumo frente a terceros, la Universidad de Piura y/o la Administración Pública toda responsabilidad que pueda derivarse por el trabajo final presentado. Lo señalado incluye responsabilidad pecuniaria incluido el pago de multas u otros por los daños y perjuicios que se ocasionen.

La asesoría del trabajo estuvo a cargo de la siguiente docente de la Universidad de Piura:

- Celia Patricia Vera Rojas, identificado con DNI: 10559658

Declaramos que:

Luego de haber empleado el software de coincidencia Turnitin, revisado las fuentes de información señaladas por el autor, y en razón de nuestra experiencia como investigadores, declaramos que las ideas expuestas en el trabajo final alcanzan las condiciones de calidad, integridad y originalidad acorde a los objetivos institucionales y estándares en materia de investigación. Finalmente, no asumimos responsabilidad por la posible vulneración de derechos de autor en el trabajo final referido, pues tal responsabilidad es exclusiva del autor.

Fecha: 24/06/2025.



.....
Firma del autor¹



.....
Firma del asesor¹

¹ Firma idéntica al DNI. No se admite digital, salvo certificado.



UNIVERSIDAD
DE PIURA

Declaración Jurada de Originalidad del Trabajo Final

Yo, Macarena Ayleen Mansilla Mahmud, egresado del Programa Académico de Economía de la Facultad de Ciencias Económicas y Empresariales de la Universidad de Piura, identificado(a) con DNI: 72970452, declaro que:

Soy autor del trabajo final titulado:

"El impacto de la malnutrición de los estudiantes en su rendimiento escolar: Un estudio en niños y adolescentes peruanos"

El mismo que presento bajo la modalidad de Tesis para optar el Título profesional de Economista.

Que el trabajo se realizó en coautoría con los siguientes alumnos de la Universidad de Piura.

- Mauricio Andre Beuzeville Montañez, identificado con DNI: 71888762

El texto de mi trabajo final es original y no vulnera los derechos de terceros o, de ser el caso, derechos de los coautores, incluidos los derechos de propiedad intelectual, datos personales, entre otros. En tal sentido, el texto de mi trabajo final no ha sido plagiado total ni parcialmente, para lo cual, he respetado las normas internacionales de citas y referencias de las fuentes consultadas. Asimismo, el texto del trabajo final que presento no ha sido publicado ni presentado antes en cualquier medio electrónico o físico; y que la investigación, los resultados, datos, conclusiones y demás información presentada que atribuyo a mi autoría son veraces.

En caso de detectarse el incumplimiento de lo declarado asumo frente a terceros, la Universidad de Piura y/o la Administración Pública toda responsabilidad que pueda derivarse por el trabajo final presentado. Lo señalado incluye responsabilidad pecuniaria incluido el pago de multas u otros por los daños y perjuicios que se ocasionen.

La asesoría del trabajo estuvo a cargo de los siguientes docentes de la Universidad de Piura:

- Celia Patricia Vera Roja, identificado con DNI: 10559658

Declaro (declaramos) que:

Luego de haber empleado el software de coincidencia Turnitin, revisado las fuentes de información señaladas por el autor, y en razón de mi (nuestra) experiencia como investigador(es), declaro (declaramos) que las ideas expuestas en el trabajo final alcanzan las condiciones de calidad, integridad y originalidad acorde a los objetivos institucionales y estándares en materia de investigación. Finalmente, no asumo (asumimos) responsabilidad por la posible vulneración de derechos de autor en el trabajo final referido, pues tal responsabilidad es exclusiva del autor.

Fecha: 24/06/2025.

Firma del autor²

Firma del asesor¹

² Firma idéntica al DNI. No se admite digital, salvo certificado.

Dedicatoria

A mis abuelos María, Jaime y Mercedes, por ser mi inspiración y ejemplo de vida. Que Dios me permita seguir sus pasos y honrar su legado.

A mis padres, Jaime y Tamara, por su amor incondicional, por creer en mí incluso en los momentos en que yo dudé, y por enseñarme el valor del esfuerzo y la perseverancia.

A mi familia UDEP, por acompañarme con cariño, aliento y amistad sincera a lo largo de este camino.

A mis amigos cercanos, por su compañía, palabras de aliento y por estar presentes en los momentos clave de este proceso.

A nuestra asesora de tesis, Patricia, por su guía, paciencia y compromiso constante.

A Mauricio, mi compañero de tesis y mejor amigo, cuya presencia, entrega y apoyo incondicional fueron fundamentales para alcanzar esta meta.

Macarena Mansilla

En primer lugar, quiero agradecer profundamente a mi familia, mis amigos y mi enamorada, por enseñarme tanto desde sus aciertos como desde sus errores; por mostrarme el orden y el desorden, la humildad y el orgullo, la seriedad y el juego. Gracias por enseñarme que no todo es completamente bueno ni completamente malo, y que en la vida se aprende también de los contrastes. También deseo agradecer a mi compañera de tesis y mejor amiga, con quien compartí un largo camino lleno de aprendizajes y desafíos para llegar a este momento; así como a nuestra asesora, por sus valiosos aportes y guía constante.

Mauricio Beuzeville

Resumen

La presente investigación analiza el impacto de la malnutrición —entendida tanto como desnutrición y sobrepeso u obesidad— sobre el rendimiento académico de niños y adolescentes en el Perú. Para ello, se utilizó la base de datos *Niños del Milenio*, centrando el análisis en la cohorte mayor, y considerando evaluaciones en matemáticas y comunicación.

Esta tesis propone que la relación entre el estado nutricional (medido a través del Índice de Masa Corporal) y el rendimiento académico (tanto matemático como comunicativo) es no lineal. Para identificar este efecto, se emplean modelos de regresión semiparamétrica con Kernel considerando efectos fijos a nivel comunitario, que permiten capturar mejor la complejidad de la relación. Se utiliza además el IMC de la madre como variable instrumental, abordando problemas de endogeneidad.

Los principales hallazgos muestran que tanto la desnutrición como el sobrepeso u obesidad afectan negativamente el rendimiento escolar de manera no lineal, intensificándose en los extremos del IMC, afectando tanto a aquellos estudiantes con desnutrición, como a quienes presentan sobrepeso u obesidad, con un efecto más marcado en las habilidades comunicativas. Estos resultados refuerzan la importancia de diseñar políticas públicas integrales que aborden los extremos de la malnutrición, como parte de una estrategia para mejorar la calidad educativa en el país.

Tabla de contenido

Introducción	10
Capítulo 1 Revisión de literatura.....	14
1.1. Evidencia Internacional	14
1.2. Evidencia Nacional.....	17
1.3. Síntesis de la literatura e hipótesis.....	18
Capítulo 2 Marco Analítico.....	19
2.1. Base de datos Niños del Milenio	19
2.2. Estadística descriptiva	21
Capítulo 3 Marco teórico	32
Capítulo 4 Metodología.....	34
4.1. Discusión de endogeneidad.....	34
4.2. Relación no lineal entre el IMC y el rendimiento académico.....	36
4.3. Modelo semiparamétrico con Kernel.....	38
4.4. Especificación del modelo	39
Capítulo 5 Resultados.....	42
5.1. Modelo base: MCO con efectos fijos.....	42
5.2. Modelo VI con efectos fijos	43
5.2.1. <i>Primera etapa con VI con efectos fijos</i>	43
5.2.2. <i>Segunda etapa con VI con efectos fijos</i>	44
5.2.3. <i>Robustez del instrumento</i>	47
5.3. Modelo semiparamétrico con Kernel	47
5.3.1. <i>Estimación con márgenes predichos</i>	49
5.3.2. <i>Robustez de modelo no lineal</i>	52
Conclusiones	55
Recomendaciones	56
Referencias.....	57
Anexos	64

Anexo 1 Registro de observaciones eliminadas en la base de datos	64
Anexo 2 Comunidades de Niños de Milenio.....	64
Anexo 3 Especificación de la función Kernel de Epanechnikov	65
Anexo 4 Resultados de análisis de subgrupos para el rendimiento matemático	66
Anexo 5 Resultados de análisis de subgrupos para el rendimiento comunicativo	67



Lista de tablas

Tabla 1 Edad de Niños del Milenio por cohortes y rondas.....	19
Tabla 2 Cantidad de entrevistados de Niños del Milenio por cohortes y rondas.....	21
Tabla 3 Rendimiento promedio de evaluaciones de matemáticas normalizados por ronda.....	22
Tabla 4 Rendimiento promedio de evaluaciones de comunicación normalizados por ronda	23
Tabla 5 Rangos de las categorías del estado nutricional de los individuos	23
Tabla 6 Porcentaje de estudiantes por categoría de IMC, según sexo y etapa de desarrollo	24
Tabla 7 Estadística descriptiva del índice de riqueza por rondas	25
Tabla 8 Distribución de estudiantes según variables de control por rondas	26
Tabla 9 Porcentaje por categoría de IMC de la madre por ronda	27
Tabla 10 Resultados del modelo de Mínimos Cuadrados Ordinarios con efectos fijos	42
Tabla 11 Resultados de la primera etapa del modelo con variable instrumental con efectos fijos	44
Tabla 12 Resultados de la segunda etapa del modelo con variable instrumental con efectos fijos.....	45
Tabla 13 Comparativa entre resultados de los modelos MCO y VI.....	46
Tabla 14 Test de Sargan-Hansen.....	47
Tabla 15 Resultados del modelo semiparamétrico del efecto del IMC sobre el rendimiento matemático y comunicativo.....	48
Tabla 16 Márgenes del IMC en el rendimiento académico matemático	49
Tabla 17 Márgenes del IMC en el rendimiento académico comunicativo.....	51

Lista de figuras

Figura 1 Departamentos con localidades estudiadas en el Perú	20
Figura 2 Gráfico de dispersión del IMC entre madres e hijos.....	27
Figura 3 Densidades de Kernel por rondas de puntajes de matemáticas (Rendimiento matemático) y de PPVT normalizados (Rendimiento comunicativo)	28
Figura 4 Gráfico de dispersión entre rendimiento matemático e IMC del individuo.....	30
Figura 5 Gráfico de dispersión entre rendimiento comunicativo e IMC del individuo	30
Figura 6 Suavizamiento polinomial entre IMC y matemáticas	50
Figura 7 Suavizamiento polinomial entre IMC y PPVT	52



Introducción

En el Perú, los resultados obtenidos respecto al rendimiento escolar son poco alentadores. En efecto, según el Programa para la Evaluación Internacional de los Estudiantes (PISA, Programme for International Student Assessment, en inglés) elaborado por la Organización para la Cooperación y el Desarrollo Económico (OCDE), si bien existen algunas mejoras en los resultados de comprensión lectora y matemáticas entre los años 2009 y 2022, todavía se evidencian rezagos si los comparamos con los registrados en países de la región como Chile y Uruguay, principalmente. Asimismo, según muestran las Evaluaciones Censales de Estudiantes (ECE) del Ministerio de Educación del Perú (Minedu), entre los años 2018 y 2019 se han identificado retrocesos en comprensión de lectura y matemáticas para los diferentes niveles educativos. Por ejemplo, en el segundo grado de educación secundaria, el porcentaje de alumnos que alcanzaron los aprendizajes de comprensión de textos disminuyó de 16.2% en 2018 a 14.5% en 2019, mostrando un declive en el rendimiento en 1.7 puntos porcentuales. De la misma manera, en matemáticas, se observó una ligera disminución en segundo grado de secundaria, pasando de 34.8% en 2018 al 34.5% en 2019. A partir de estos resultados, resulta de suma importancia analizar e identificar qué factores podrían estar influyendo en los bajos niveles del rendimiento escolar, siendo esta una base importante para su desarrollo personal y profesional a largo plazo.

Por otro lado, en el Perú se registran altas tasas de malnutrición, manifestadas en forma de desnutrición crónica como de sobrepeso y obesidad. Según datos del Instituto Nacional de Estadística e Informática (INEI) en 2020, aproximadamente el 57% de la población peruana presenta sobrepeso u obesidad, lo que representa una preocupante tendencia hacia problemas de salud. Asimismo, en el 2021, alrededor del 11.4% de los adolescentes en Perú padecían de desnutrición crónica, mientras que el sobrepeso y la obesidad afectaban al 8.6% de niños menores a 5 años (UNICEF, 2022). A su vez, se ha estimado que 3 de cada 10 adolescentes mujeres entre los 15 y 19 años presentan sobrepeso u obesidad y, al llegar a la adultez, más de la mitad vivirán con esta condición (INEI, 2020). Según un análisis de tendencias, las cifras están aumentando rápidamente, proyectándose que para el 2030 más de 1 millón de niños y adolescentes peruanos, entre 5 y 19 años, tendrán obesidad (Lobstein et al, 2019). Esta situación podría afectar tanto la salud física como mental de los jóvenes, influyendo en su rendimiento académico de diversas formas.

En consecuencia, la malnutrición trae consigo diversos efectos que no solo impactan en la salud de la población sino también en su desempeño académico. Al respecto, la presente investigación se centrará en analizar las consecuencias del desempeño académico vinculadas a la desnutrición y sobrepeso u obesidad. En relación con la desnutrición, esta se refiere a una condición en la que una persona no recibe suficientes nutrientes esenciales durante un período prolongado, lo cual puede

deberse a una dieta deficiente en calorías, proteínas, vitaminas y minerales; y puede resultar en un crecimiento y desarrollo deficientes, tanto física como mental. Por otra parte, el sobrepeso u obesidad se define como una condición en la que se acumula una cantidad excesiva de grasa en el organismo, lo que provoca un aumento considerable del peso corporal. Ambas afecciones suelen evaluarse mediante el Índice de Masa Corporal (en adelante, IMC), el cual relaciona el peso con la estatura de una persona.

Dado lo alarmante de las estadísticas respecto a los problemas de salud, es necesario ahondar en el estudio de la realidad de la malnutrición y cómo esta puede impactar en el ámbito educativo. En este sentido, existen diversas investigaciones que evidencian que las altas tasas de malnutrición en el Perú, que abarca desde la desnutrición crónica hasta el sobrepeso y la obesidad, representan una grave amenaza para la educación (Benavides, 2002; Cueto, 2005; Mendoza, 2019). Asimismo, se debe destacar que esta problemática es de relevancia económica, dado que el rendimiento académico escolar impacta al desarrollo individual (Chetty et al., 2011), así como, las habilidades cognitivas de la población se correlacionan con el crecimiento económico del país (Hanushek y Woessmann, 2008).

La relación entre la malnutrición y el rendimiento escolar es un elemento importante a considerar en las políticas públicas. Estos problemas de salud se manifiestan de distintas maneras, desde deficiencias nutricionales que afectan el desarrollo cognitivo (Hioui et al., 2016) hasta el impacto negativo en la salud, no solo física, sino también mental de los estudiantes (Ramsey y Muskin, 2013; Chinyoka, 2014). En particular, la malnutrición puede incidir directamente en la capacidad de concentración y aprendizaje de una persona en lo académico, dada la falta de nutrientes esenciales y vitaminas, lo que puede traducirse en diversas consecuencias como la disminución de la atención y alteraciones en la función cognitiva, los cuales son variables que inciden de manera significativa en el desempeño académico (Richardson & Ross, 2000).

Además, esta conexión entre la salud y el rendimiento académico no solo ocurre en términos individuales, sino también a nivel agregado; es decir, cuando un grupo de estudiantes de un país enfrenta esta misma problemática, el país ve seriamente afectada su capacidad para progresar hacia un desarrollo educativo y académico más alto (Themane et al., 2003; Zaini et al., 2005). Por lo tanto, la comprensión de la importancia de la relación entre la malnutrición y el rendimiento académico es un esfuerzo estratégico para ayudar a abordar los diferentes problemas detrás del bajo rendimiento educativo en Perú. Considerando el contexto detallado, la presente investigación tiene como motivación analizar cómo la malnutrición, medida a través del IMC, afecta al rendimiento académico, tanto comunicativo como matemático, y determinar si esta relación es lineal.

El objetivo de este estudio es identificar el impacto de la malnutrición de los estudiantes peruanos en su rendimiento escolar. Para ello, empleamos inicialmente un modelo de regresión lineal que nos permite medir el efecto de la malnutrición, tanto de la desnutrición como de la sobrepeso u

obesidad, en el rendimiento académico. Sin embargo, la literatura sugiere que esta relación presenta problemas de endogeneidad, los cuales abordamos usando el modelo de mínimos cuadrados ordinarios en dos etapas utilizando efectos fijos con una variable instrumental, la cual será el IMC de la madre, debido a la carga genética que esta aporta en los hijos.

En el contexto peruano, la investigación sobre esta relación es limitada, destacando Saintila y Rodríguez (2016), Monge, Campana y Gutiérrez (2017), Cueto (2005), Rojas et al. (2015), entre otros. De estos, Cueto (2005) desarrolló un estudio con niños de veinte escuelas de dos zonas rurales del Perú, sin embargo, no encontró asociación entre la malnutrición -aproximada por variables antropométricas-, con el rendimiento en matemáticas y comprensión lectora. No obstante, halló una asociación positiva entre el IMC y la deserción escolar, específicamente, los niños con un IMC más alto en 1998 tenían más probabilidades de haber abandonado la escuela para 2001, lo cual sugiere que los niños con un IMC mayor eran más propensos a dejar de asistir a clases, planteando que esto podría deberse a expectativas culturales.

Por otro lado, en la literatura nacional no existen estudios que consideren la no linealidad del IMC en el rendimiento académico, ya que han empleado modelos lineales que no capturan adecuadamente las posibles variaciones en esta relación. Es por ello que, en la presente investigación, utilizamos modelos semiparamétricos, los cuales permiten analizar el vínculo entre el IMC y el rendimiento académico sin imponer una forma funcional rígida (Horowitz, 2009), lo que permitirá observar cómo varía esta relación en distintos subgrupos o rangos del IMC (Shimokawa, 2008).

Para el desarrollo de dichas metodologías, empleamos la base de datos de Niños de Milenio, específicamente la cohorte mayor, efecto que se medirá tomando en consideración las características individuales (controles). Asimismo, esta base nos permite ampliar la cobertura geográfica, lo cual brindará una visión más completa sobre esta relación. Un aporte adicional del estudio es la inclusión del control índice de riqueza, el cual permite identificar el efecto de contar con mayor bienes y servicios; así como la inclusión de la variable del tipo de colegio, sea público o privado.

La hipótesis de la presente investigación plantea que existe una relación no lineal entre el peso corporal, medido a través del IMC, y el rendimiento académico, tanto matemático como comunicativo. Esta relación variaría según los diferentes niveles del IMC, siendo positiva cuando el IMC se encuentra en rangos normales y tornándose negativa en los valores que se alejan de este rango óptimo (tanto bajos como altos), manteniendo una tendencia general negativa.

Los resultados principales muestran que el IMC tiene un efecto negativo en el rendimiento académico, siendo este impacto más pronunciado en el rendimiento comunicativo que en el matemático. Asimismo, se encuentra que esta relación es no lineal, donde tanto los valores extremos bajos (desnutrición) como los valores extremos altos (sobrepeso-obesidad) tienen efectos negativos en el rendimiento académico, a diferencia del IMC en rangos normales.

Este estudio está estructurado de la siguiente manera: el Capítulo 1 presenta una revisión de literatura sobre las variables que impactan en el rendimiento académico, con énfasis en la malnutrición, en el Capítulo 2 se describe el marco analítico, en el Capítulo 3 se presenta el marco teórico, el Capítulo 4 informa sobre la metodología empleada, el Capítulo 5 incluye los resultados obtenidos y, finalmente, presentamos las conclusiones y recomendaciones.



Capítulo 1 Revisión de literatura

En los últimos años, tanto a nivel internacional como nacional, se ha observado una creciente preocupación por los efectos de la desnutrición y el sobrepeso/obesidad en el rendimiento académico de los niños y adolescentes. Este fenómeno no solo representa un problema de salud pública, sino que también tiene implicaciones en el desarrollo cognitivo y el futuro laboral de los jóvenes.

1.1. Evidencia Internacional

A nivel internacional, diversos estudios han explorado esta relación. Así, por ejemplo, Datar et al. (2004), Crosnoe y Muller (2004), y Averett y Stifel (2010) concluyen que la obesidad afecta negativamente a los escolares en sus puntajes de las pruebas cognitivas. De igual forma, Must y Strauss (1999) postulan que la obesidad infantil está relacionada con problemas de salud que interfieren en el aprendizaje, tales como la apnea del sueño, la hipertensión y la diabetes. Currie y Stabile (2006), por su parte, sostienen que la obesidad también causa ausencia escolar, estrés y fatiga que pueden interferir negativamente en el rendimiento escolar.

Además, investigaciones neurobiológicas, como las de Gunstad et al. (2008), han encontrado que un mayor IMC se correlaciona con una disminución en el volumen de materia gris en el cerebro, lo que está relacionado con un peor rendimiento cognitivo. Maureira et al. (2019) sugieren una correlación negativa entre el IMC y el rendimiento académico, indicando que los estudiantes con un IMC más alto tienden a tener calificaciones académicas ligeramente más bajas en diversas escuelas de Santiago de Chile. Es importante mencionar que estas investigaciones fueron de naturaleza correlacional y no establecen relaciones causales.

Por otro lado, los niños y adolescentes con sobrepeso u obesidad experimentan dificultades en la escuela, y se ha demostrado que los déficits de memoria de trabajo explican parcialmente el bajo rendimiento académico de los niños obesos (Gunnarsdottir, et. al.,2015; Sánchez-Cruzat et al.,2017). Respecto a los resultados diferenciados por género, Sabia (2007) observó que el peso corporal tiene un impacto más significativo en el rendimiento académico de las mujeres, particularmente en las mujeres blancas, donde se encontró una relación negativa consistente entre el IMC y el promedio de calificaciones. En este grupo, tanto el sobrepeso como la percepción de ser "con sobrepeso" están asociados con una disminución en el rendimiento académico. En contraste, para los hombres, la relación entre el peso y el rendimiento académico fue menos consistente, y en el caso de los hombres blancos, no se detectó una relación significativa. A su vez, Datar et. al. (2004), y Ding et al. (2009) encontraron resultados similares, en donde las niñas con sobrepeso tienden a tener peores calificaciones en comparación con los niños.

Algunos estudios han empleado variables instrumentales para abordar la posible endogeneidad, los cuales proporcionan evidencia robusta sobre la relación causal entre el IMC y el rendimiento académico. Destaca Von Hinke et al. (2012), en el cual se identifican tres conjuntos diferentes de instrumentos que se han empleado en esta relación: a) IMC del niño en periodos anteriores como instrumento para el IMC actual (Kaestner y Grossman, 2009); b) dos marcadores genéticos previamente utilizados como instrumentos para el IMC actual (Von Hinke et al., 2016); y c) IMC de la madre antes del embarazo como instrumento para el IMC del niño (como en Sabia, 2007; Averett y Stifel, 2010).

Kaestner y Grossman (2009) analizan a niños de 5 a 12 años entre 1986 y 2004, regresionando el cambio en el logro educativo a lo largo de dos años sobre indicadores que representan el estado de bajo peso y sobrepeso del niño, y utilizan variables instrumentales, especificando los percentiles de IMC rezagados del niño como instrumentos para el peso actual. Von Hinke et al. (2016) presentan una discusión sobre las condiciones que deben cumplirse para que los marcadores genéticos puedan ser utilizados como instrumentos, examinando los efectos de la masa grasa de los niños en su rendimiento académico. Emplean las variantes genéticas FTO y MC4R como variables instrumentales para la masa grasa medida mediante un escaneo DXA. Sin embargo, en ambas investigaciones no se encuentra evidencia de que el progreso académico de los niños se vea afectado por su peso.

No obstante, en Sabia (2007) se analizan a adolescentes de 14 a 17 años y encuentra una relación negativa entre el IMC de las niñas blancas y su rendimiento educativo. Estos resultados son robustos tanto a un enfoque de efectos fijos individuales como a una especificación que utiliza variables instrumentales, empleando como instrumentos el estado de obesidad autoinformado de la madre y el padre. Asimismo, en Averett y Stifel (2010) se centran en niños de primaria (6–13 años), mostrando que aquellos con sobrepeso tienen resultados educativos más bajos en comparación con los niños con un peso saludable. Este hallazgo se mantiene al utilizar tanto efectos fijos individuales como variables instrumentales, empleando como instrumentos el IMC de la madre antes del embarazo y el cuadrado de la misma variable.

Es importante señalar que varios estudios han evidenciado que la obesidad puede tener un impacto psicosocial que, a su vez, afecta el desempeño académico. Puhl y Latner (2007) identificaron que el acoso escolar y la ansiedad, comunes entre los adolescentes con sobrepeso, contribuyen a una baja autoestima y menor motivación para el aprendizaje. Mientras que Sabia et. al. (2015), encontró que, en Estados Unidos, el peso corporal tiene efectos negativos en la autoestima y la sintomatología depresiva en mujeres, aunque estos efectos no se manifestaron en el caso de los hombres. Además, el estudio identificó una relación desfavorable entre el peso corporal y el rendimiento académico en mujeres, y señaló que el bienestar psicosocial pudo explicar hasta un 30% de esta relación.

No obstante, en estos casos puede surgir un problema de endogeneidad, el cual ha sido abordado en diversas investigaciones con metodologías diferentes, como las de Neumark-Sztainer et al. (2002) y Krukowski et al. (2009). Neumark-Sztainer señala que el sobrepeso en adolescentes está relacionado con mayores niveles de acoso escolar. Al controlar este factor y resolver el problema de variable omitida, encuentra que es el acoso, y no la obesidad o el sobrepeso, lo que explica el menor rendimiento académico. Krukowski, por su parte, va más allá al introducir el concepto de "*weight based teasing*" o acoso específico por sobrepeso. Al controlar esta variable, concluye que este factor es el que explica el bajo rendimiento académico, sugiriendo que el efecto es más de índole social que físico.

En cuanto a la desnutrición, Vashti et. al. (2018) afirma que existe una relación significativa entre la desnutrición en la etapa de la niñez y el nivel educativo, lo que indica que la desnutrición puede afectar negativamente el rendimiento académico. Selamawit et. al. (2021) analiza la asociación entre el estado nutricional y el rendimiento académico entre los estudiantes adolescentes, identificando que la desnutrición puede inhibir el rendimiento académico debido a un crecimiento deficiente, un retraso mental y una función cognitiva defectuosa. La deficiencia de nutrientes esenciales, como el hierro y el yodo, afecta la capacidad de concentración y aprendizaje, lo que compromete el desempeño académico. Chang et al. (2010) demostraron que estudiantes con historial de desnutrición crónica obtuvieron puntajes significativamente más bajos en pruebas estandarizadas de matemáticas y lectura, incluso después de controlar por factores socioeconómicos.

Asimismo, Hoddinott et al. (2013) encontraron que los niños que experimentaron desnutrición antes de los dos años mostraron un rendimiento académico significativamente menor en pruebas de lectura y matemáticas durante su etapa escolar; mientras que Crookston et al. (2013) evidenciaron que la recuperación del retraso en el crecimiento debido a la desnutrición está asociada con mejoras en el rendimiento cognitivo y académico, reforzando la importancia de la existencia de programas escolares que buscan reducir la desnutrición infantil severa.

Finalmente, en Victora et al. (2008) encontraron que la desnutrición tiene efectos acumulativos: los déficits nutricionales tempranos no solo afectan el rendimiento académico inmediato, sino que tienen consecuencias a largo plazo en el desarrollo cognitivo y el logro educativo; mientras que, en Kar et al. (2008) identificaron que los niños con desnutrición muestran déficits específicos en atención, concentración, memoria de trabajo y velocidad de procesamiento cognitivo, lo que afecta directamente su capacidad de aprendizaje y rendimiento escolar.

Estos estudios proporcionan evidencia sólida sobre los efectos negativos a largo plazo de la malnutrición en el rendimiento académico y resaltan la importancia de la intervención temprana.

1.2. Evidencia Nacional

En el contexto peruano, la investigación sobre la relación entre el estado nutricional y el rendimiento académico es limitada, pero se ha encontrado evidencias relevantes. Cueto (2005) desarrolló un estudio con niños de veinte escuelas en dos zonas rurales del Perú, sin embargo, no encontró asociación entre la malnutrición -aproximada por variables antropométricas-, con el rendimiento en matemáticas y comprensión lectora. No obstante, halló una asociación positiva entre el IMC y la deserción escolar, específicamente, los niños con un IMC más alto en 1998 tenían más probabilidades de haber abandonado la escuela para 2001, lo cual sugiere que los niños relativamente con mayor peso (con un IMC mayor) eran más propensos a dejar de asistir a clases, planteando que esto podría deberse a expectativas culturales.

Por otro lado, Saintila y Rodríguez (2016) reportaron que la desnutrición tiene un impacto negativo en el rendimiento académico de niños de 7 a 14 años en una escuela de Lima. Del mismo modo, un estudio más reciente de Monge, Campana y Gutiérrez (2017), que abarcó una muestra nacional de 2,415 observaciones, concluyó que la desnutrición crónica tiene un efecto negativo significativo en el rendimiento académico de los niños peruanos. Asimismo, en un estudio exploratorio en la región amazónica de Iquitos, Rojas et al. (2015) encontraron una asociación positiva entre el estado nutricional y el desempeño académico, aunque es importante señalar que este estudio fue de naturaleza correlacional y no establece relaciones causales. Por su parte, Wisniewski (2017) utilizó tres rondas de la base de datos de Niños del Milenio y, tras corregir por posibles sesgos derivados de heterogeneidad no observable, identificó una relación débilmente significativa entre la obesidad en niñas y un mayor rendimiento. Por otro lado, identificó que niños indígenas con desnutrición tienen puntuaciones cognitivas significativamente más bajas, relación que se mantiene luego de corregir por posible sesgo de heterogeneidad.

En resumen, a pesar de que la relación entre el estado nutricional y el rendimiento académico ha sido explorada en algunos estudios en el contexto peruano, la evidencia sigue siendo limitada y presenta varias carencias. Investigaciones previas, como las de Cueto (2005) y Monge et al. (2017), han encontrado efectos negativos de la desnutrición en el rendimiento académico, pero estos estudios no han abordado de manera exhaustiva la complejidad de la relación entre el IMC y el rendimiento escolar. Además, se han centrado en muestras locales o en enfoques que no permiten una comprensión completa de la variabilidad en el impacto del estado nutricional a lo largo de las diferentes etapas del desarrollo escolar.

La presente investigación se enfoca en superar estas limitaciones al emplear modelos semiparamétricos que permiten analizar las relaciones entre el IMC y el rendimiento académico sin imponer una forma funcional rígida. Además, se busca ampliar la cobertura geográfica utilizando la base de datos de Niños del Milenio, lo que permitirá obtener una visión más completa de la relación

entre malnutrición y rendimiento académico a nivel nacional. Al controlar variables a nivel individual, esta investigación ofrece una perspectiva más integral sobre los factores que afectan el desempeño académico de los estudiantes peruanos, lo que constituye una contribución significativa a la literatura existente.

1.3. Síntesis de la literatura e hipótesis

La relación entre la malnutrición (desnutrición y sobrepeso/obesidad) y el rendimiento académico en niños y adolescentes es un tema de creciente preocupación tanto a nivel internacional como nacional. Diversos estudios internacionales han encontrado que la obesidad afecta negativamente el desarrollo cognitivo y el rendimiento académico debido a factores como problemas de salud (apnea del sueño, hipertensión), ausencias escolares y estrés. Investigaciones neurobiológicas sugieren que un mayor IMC está relacionado con una disminución del volumen de materia gris, lo que compromete el rendimiento cognitivo. Además, el acoso escolar y la ansiedad relacionados con el sobrepeso también influyen en el bajo rendimiento académico, especialmente en mujeres, según algunos estudios que destacan la importancia del bienestar psicosocial.

Por otro lado, la desnutrición se asocia con dificultades en el crecimiento, retraso mental y deficiencias cognitivas que limitan el rendimiento académico. Investigaciones internacionales muestran que la deficiencia de nutrientes esenciales, impacta en la concentración y el aprendizaje, afectando negativamente el desempeño escolar. Estos estudios destacan la necesidad de implementar estrategias nutricionales que aborden ambos extremos de la malnutrición para mejorar el rendimiento académico y reducir la pobreza a largo plazo.

La hipótesis base de la presente investigación, consistente con la evidencia específica existente en los estudios detallados sobre el rendimiento académico y la malnutrición, plantea que existe una relación no lineal entre el peso corporal, medido a través del IMC, y el rendimiento académico, tanto matemático como comunicativo. Esta relación varía según los diferentes niveles de IMC, siendo positiva cuando el IMC se encuentra en rangos normales y tornándose negativa en los valores que se alejan de este rango óptimo, manteniendo una tendencia general negativa del IMC en el rendimiento.

Capítulo 2 Marco Analítico

2.1. Base de datos Niños del Milenio

El estudio "Niños del Milenio" o "Young Lives" es una investigación sobre la pobreza infantil que ha seguido el desarrollo de 12,000 niños en Etiopía, India, Perú y Vietnam a lo largo de 15 años. Desde 2002, Young Lives ha estado monitoreando dos cohortes de niños, una nacida en 1994 y otra en 2001. El propósito del estudio es generar información que permita comprender las causas y efectos de la pobreza en la infancia, y analizar cómo las políticas públicas, implementadas o por implementar, impactan el bienestar infantil.

En el caso de Perú, la muestra original fue seleccionada al azar en 20 localidades de 14 departamentos del país, aunque el 5% de los distritos más ricos se excluyeron del muestreo. El estudio sigue la vida de dos grupos de niños durante quince años. El primer grupo, conformado por unos 2,000 niños, tenía entre 6 y 18 meses (cohorte menor) en 2002, cuando se recogieron los primeros datos. El segundo grupo, que será el foco de este estudio, incluye a aproximadamente 750 niños que tenían entre 7 y 8 años (cohorte mayor) en 2002. La investigación se lleva a cabo a través de encuestas y entrevistas realizadas a los niños, sus padres y representantes de sus comunidades, seleccionados aleatoriamente entre 100 niños de los 14 departamentos del Perú. En la siguiente tabla se presentan las edades de las cohortes por rondas:

Tabla 1

Edad de Niños del Milenio por cohortes y rondas

Cohorte	Ronda 1 2002	Ronda 2 2006	Ronda 3 2009	Ronda 4 2012	Ronda 5 2016
Menor	6-8 meses	4-5 años	7-8 años	11-12 años	14-15 años
Mayor	7-8 años	11-12 años	14-15 años	18-19 años	21-22 años

Adaptado de: Datos de Niños del Milenio, rondas 1-5. Elaboración propia.

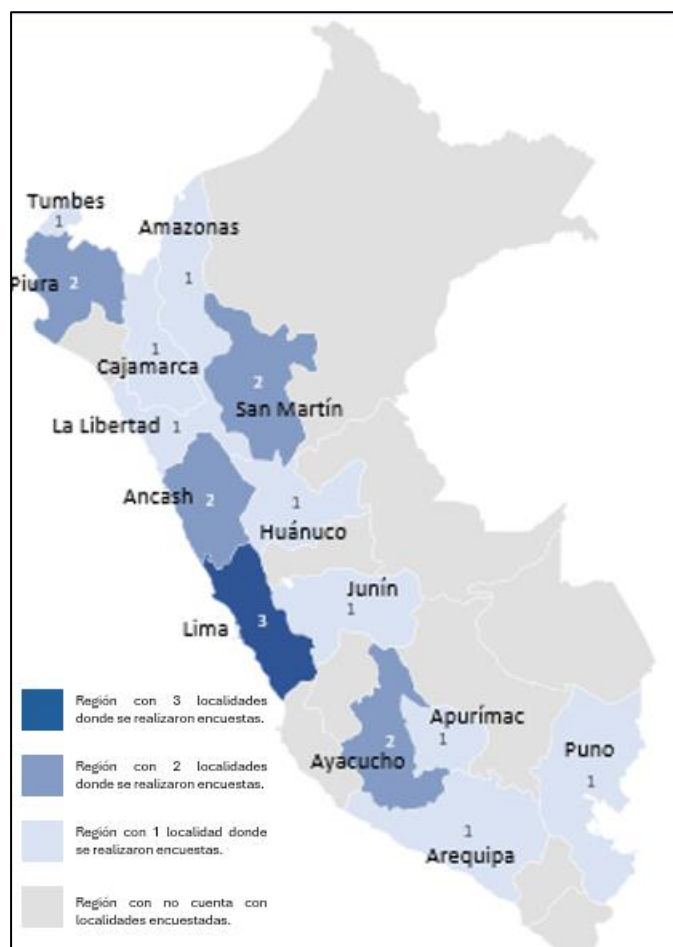
Como se mencionó anteriormente, se realizan encuestas de hogares a gran escala dirigidas a niños y sus cuidadores principales, las cuales se complementaron con métodos cualitativos como entrevistas detalladas, sesiones grupales y análisis de casos específicos; involucrando a una muestra seleccionada de niños, padres o tutores, educadores y líderes locales. El estudio recopiló datos extensos que abarcan no solo las condiciones materiales y sociales de los participantes, sino también sus propias percepciones sobre su vida actual y sus expectativas. Este enfoque integral considera el contexto ambiental y sociocultural de las comunidades estudiadas.

Siguiendo estos principios generales, el equipo en Perú implementó una metodología de muestreo que presentó variaciones significativas en comparación con los enfoques utilizados en los otros tres países participantes. Mientras que en estos últimos se optó por una estrategia de selección

de conglomerados que combinaba elementos aleatorios e intencionales, el estudio peruano se distinguió por adoptar un enfoque aleatorio en su proceso de muestreo. En la siguiente figura, se destacan los departamentos del Perú en donde se han llevado a cabo las encuestas de Niños del Milenio:

Figura 1

Departamentos con localidades estudiadas en el Perú



Adaptado de: Datos de Niños del Milenio. Elaboración propia.

Dentro de las limitaciones de Niños del Milenio se encuentra la pérdida total de la muestra a lo largo de los años. La investigación empezó con aproximadamente 2000 niños para la cohorte menor y 750 para la cohorte mayor. No obstante, para la tercera ronda, la pérdida total de la muestra fue de 4.4% durante el período de ocho años, mientras que para la ronda 5 se aumentó esta pérdida total. En la siguiente tabla se presenta la cantidad de entrevistados por rondas:

Tabla 2*Cantidad de entrevistados de Niños del Milenio por cohortes y rondas*

Cohorte	Ronda 1 2002	Ronda 2 2006	Ronda 3 2009	Ronda 4 2012	Ronda 5 2016
Menor	-	2052	1943	1902	1860
Mayor	714	714	678	635	618

Adaptado de: Datos de Niños del Milenio, rondas 1-5. Elaboración propia.

Para esta investigación se ha decidido utilizar las rondas 2 a la 4 de la cohorte mayor, ya que los datos necesarios se encuentran disponibles en estas fases del estudio. Durante el proceso de limpieza de datos, se eliminaron observaciones a nivel del niño en las cuales había valores faltantes consideradas importantes para el análisis; así como se eliminaron entradas duplicadas que podrían haber sesgado los resultados. En cuanto a la detección de *outliers*, se identificaron valores atípicos en la variable del rendimiento académico (matemático y comunicativo) y del IMC, tanto del individuo como de la madre, que no tenían sentido o parecían erróneos. Como resultado del proceso de limpieza, se eliminaron un total de 290 observaciones, incluyendo valores faltantes, datos duplicados y valores atípicos (Anexo 1).

Adicionalmente, se realizó la estandarización de formatos para asegurar que los datos numéricos, fechas y textos se encontraran en un formato consistente. Por último, se normalizaron los puntajes de las pruebas de vocabulario (PPVT, por sus siglas en inglés) y matemáticas, alineando las escalas de las variables numéricas para garantizar que sean comparables en su interpretación, lo cual es fundamental para modelos que son sensibles a la magnitud de los valores. Luego de la limpieza de datos, se cuentan con 579 observaciones a nivel del individuo para cada ronda, lo que da un total de 1,737 observaciones.

2.2. Estadística descriptiva

Variable dependiente: rendimiento académico escolar

Siendo el rendimiento académico nuestra variable de interés, esta será subdividida en rendimiento matemático y comunicativo. Es importante destacar que, dada la naturaleza de la base de datos de Niños del Milenio, la cual constituye un estudio no directamente vinculado al colegio, las evaluaciones de rendimiento siguen un marco de referencia normativo. Esto implica que los puntajes se interpretan en relación con las distribuciones correspondientes; es decir, un puntaje se clasifica dentro de una o más distribuciones de puntajes o se compara con el rendimiento promedio de los examinados en diversas poblaciones de referencia (Cueto et al. 2009). Por esta razón, hemos optado por normalizar ambas variables de rendimiento, ajustándolas para que adquieran valores comprendidos entre 0 y 1. De esta manera, facilitamos su comparación en nuestros resultados.

En el caso del rendimiento matemático, la prueba realizada constó de diez elementos seleccionados del Estudio de Tendencias Internacionales en Matemáticas y Ciencias (TIMSS, por sus siglas en inglés), desarrollado por la Asociación Internacional para la Evaluación del Rendimiento Educativo (IEA, por sus siglas en inglés) en 2003. Se seleccionaron elementos con diferentes niveles de dificultad con el objetivo de discriminar entre los estudiantes de bajo y alto rendimiento. La prueba tuvo un límite de tiempo estricto, pero se instruyó a los entrevistadores a animar a los niños a completarla en quince minutos. Si bien a lo largo de las rondas las evaluaciones matemáticas incorporaron modificaciones¹, estas se fueron construyeron conservando algunos elementos para permitir la vinculación de los puntajes de rendimiento a lo largo del tiempo.

En la Tabla 3, se presentan los rendimientos promedio de las evaluaciones de los estudiantes en matemáticas normalizados entre 0 y 1, desglosados por ronda y categorizados según el sexo de los estudiantes. Se puede observar que, el rendimiento promedio es creciente a lo largo de las rondas para ambos sexos.

Tabla 3

Rendimiento promedio de evaluaciones de matemáticas normalizados por ronda

Sexo	Infancia	Adolescencia	
	Ronda 2	Ronda 3	Ronda 4
Masculino	0.478	0.53	0.591
Femenino	0.499	0.505	0.56
Total	0.488	0.517	0.575

Adaptado de: Datos de Niños del Milenio, rondas 2-4. Elaboración propia.

En el caso del rendimiento comunicativo, se evalúa a través de la Prueba de Vocabulario Peabody con Imágenes, conocida como PPVT por sus siglas en inglés, la cual es una prueba de vocabulario receptivo, es decir, de adquisición de vocabulario, diseñada para niños, adolescentes y adultos. El PPVT consta de diferentes elementos de vocabulario que se presentan en orden de dificultad creciente. Cada elemento tiene cuatro figuras simples dispuestas en un formato de opción múltiple. Se le pide a la persona que está siendo examinada que seleccione o señale la imagen que mejor ilustra el significado de una palabra presentada oralmente por el examinador. En el caso de la base de datos de Niños del Milenio en Perú se administró la versión en español llamada Prueba de Vocabulario Peabody con Imágenes Revisada o PPVT-R por sus siglas en inglés, la cual está compuesta por 125 elementos a evaluar. Dicha evaluación produjo puntajes estándar, los cuales han sido transformados utilizando una tabla de conversión en el manual del instrumento con el fin de generar puntajes con niveles aceptables de confiabilidad y validez. De manera similar a la prueba de

¹ Considerando la interpretación de datos, resolución de problemas numéricos, medición, y geometría básica.

matemáticas, a lo largo de las rondas, basándose en la prueba base, se fueron considerando algunas modificaciones debido a las traducciones presentes en la versión de PPVT-R con respecto a la versión original.

En la Tabla 4, se presentan los rendimientos promedio de las evaluaciones de los estudiantes en comunicación normalizados entre 0 y 1, desglosados por ronda y categorizados según el sexo de los estudiantes. Se puede observar que, a diferencia del rendimiento promedio matemático, este es creciente en la ronda 3, mientras que para la ronda 4 hay una caída en ambos sexos.

Tabla 4

Rendimiento promedio de evaluaciones de comunicación normalizados por ronda

Sexo	Infancia	Adolescencia	
	Ronda 2	Ronda 3	Ronda 4
Masculino	0.715	0.845	0.592
Femenino	0.804	0.854	0.533
Total	0.759	0.849	0.562

Adaptado de: Datos de Niños del Milenio, rondas 2-4. Elaboración propia.

Variables independientes: IMC y variables de control

Según la Organización Mundial de la Salud (OMS): *“la obesidad y el sobrepeso son definidas como la acumulación anormal o excesiva de grasa que puede ser perjudicial para la salud, y que es provocada por un desequilibrio entre calorías consumidas y gastadas, donde las primeras son mayores que las segundas; mientras que define a la desnutrición como una condición que resulta de una ingesta inadecuada de nutrientes o una mala absorción de estos, lo que puede generar una serie de trastornos en el cuerpo”*. Usualmente se utiliza el IMC para evaluar el estado nutricional de las personas, donde los rangos de las categorías del estado nutricional de los individuos están presentados en la Tabla 5 a continuación:

Tabla 5

Rangos de las categorías del estado nutricional de los individuos

Categorías	Rangos
Desnutrición	$IMC < 18.5$
Normal	$18.5 \leq IMC < 25$
Sobrepeso	$25 \leq IMC < 30$
Obesidad	$IMC \geq 30$

Adaptado de: OMS. Elaboración propia.

A continuación, en la Tabla 6, se presenta una breve descripción estadística de los porcentajes de estudiantes en cada categoría del IMC diferenciadas por sexo y rondas, las cuales abarcan ronda 2, 3 y 4. Se observa que en las rondas 2 y 3 la mayoría de los estudiantes, tanto masculinos como femeninos, presentan un mayor porcentaje dentro de la categoría del IMC catalogada como normal. Sin embargo, esta tendencia difiere en la ronda 4, donde el mayor porcentaje para el sexo femenino se encuentra en sobrepeso/obesidad².

Tabla 6

Porcentaje de estudiantes por categoría de IMC, según sexo y etapa de desarrollo

Etapa de desarrollo	Categorías por IMC	Sexo del estudiante	
		Masculino	Femenino
Ronda 2			
Infancia	Desnutrición	13.5	13.7
	Normal	56.3	58.4
	Sobrepeso / Obesidad	30.2	27.9
Ronda 3			
Adolescencia	Desnutrición	17.4	7.6
	Normal	52.4	56.9
	Sobrepeso / Obesidad	30.2	35.5
Ronda 4			
Adolescencia	Desnutrición	19.3	4.6
	Normal	50.8	46.6
	Sobrepeso / Obesidad	29.9	48.9

Nota. Los porcentajes están calculados sobre el total de cada sexo por ronda. Las categorías de IMC siguen los criterios de la OMS: Desnutrición ($IMC < 18.5$), Normal ($18.5 \leq IMC < 25$), Sobrepeso/Obesidad ($IMC \geq 25$). Los datos corresponden a la cohorte mayor del estudio.

Adaptado de: Datos de Niños del Milenio, rondas 2-4. Elaboración propia.

En el caso de los estudiantes de sexo masculino, el porcentaje de sobrepeso/obesidad se ha mantenido relativamente constante a lo largo de las rondas, con valores de 30.2% en la ronda 2 y 29.9% en la ronda 4, lo que indica una variación mínima. Por el contrario, en el caso de las estudiantes de sexo femenino, se registra un aumento significativo en el porcentaje de sobrepeso/obesidad durante el mismo período. Este porcentaje se incrementa notablemente del 27.8% en la ronda 2 al 48.8% en la ronda 4.

Para el caso del índice de riqueza, se incluyen variables que muestran indicativos de poder adquisitivo que varían sustancialmente en la muestra. Cabe considerar que la construcción del índice de riqueza sigue la metodología empleada en estudios previos sobre rendimiento escolar en países en

² En la presente investigación se está considerando en una sola categoría el sobrepeso con la obesidad por la escasa cantidad de datos en la base empleada respecto a estas categorías.

desarrollo (Filmer y Pritchett, 2001; Vyas y Kumaranayake, 2006), por lo cual se construye a partir de tres índices: calidad de la vivienda, el acceso a los servicios y la propiedad de bienes de consumo duraderos. Partiendo de la base en la que sus tres indicadores tienen la misma importancia, el índice de riqueza se calcula como una media simple de los tres índices como se muestra en la siguiente ecuación. Cabe resaltar que dicho resultado toma valores entre 0 y 1, donde un índice de riqueza más alto indica un estatus socioeconómico más elevado.

$$\text{Índice de riqueza} = \frac{\text{Vivienda} + \text{Servicios} + \text{Bienes}}{3} \quad [1]$$

Donde:

- *Vivienda*: Es el subíndice de calidad de la vivienda del hogar del individuo.
- *Servicios*: Es el subíndice de acceso a los servicios del hogar del individuo.
- *Bienes*: Es el subíndice de propiedad de bienes de consumo duraderos del hogar del individuo.

Para el caso del subíndice de Vivienda, refleja el bienestar de los miembros del hogar en términos de confort relacionado con la vivienda, teniendo en cuenta los materiales utilizados en la construcción de la vivienda de la familia (paredes, techo y suelo) y la densidad del hogar. Por otro lado, del subíndice de Servicios, es una medida de la capacidad del hogar para satisfacer los requisitos funcionales de una vivienda sólida. Este subíndice es una media simple de cuatro indicadores: acceso a electricidad; acceso al agua potable; acceso a un servicio de saneamiento gestionado de forma segura; y acceso a un combustible adecuado para cocina. Finalmente, para el subíndice Bienes, el subíndice de bienes de consumo duraderos es una medida de la posesión de artículos domésticos comunes por parte de los hogares. Cabe resaltar que cada indicador se calcula como la media simple de los subindicadores respectivos. En el índice de riqueza que se muestra en la Tabla 7 se puede observar una tendencia decreciente a lo largo de las rondas, pasando de un promedio de 0.496 en la ronda 2 a 0.432 en la ronda 4.

Tabla 7

Estadística descriptiva del índice de riqueza por rondas

Ronda	Índice de riqueza					Coeficiente de variación
	Media	N	Min	Max	Desviación estándar	
2	0.496	665	0.06	0.92	0.223	0.449
3	0.424	643	0.03	0.9	0.251	0.302
4	0.432	613	0.03	0.91	0.355	0.475

Adaptado de: Datos de Niños del Milenio, rondas 2-4. Elaboración propia.

En la Tabla 8, se presenta la distribución de estudiantes según las demás variables de control a lo largo de las diferentes rondas del estudio. Los datos presentan que la distribución geográfica de los estudiantes se ha mantenido relativamente estable, con una mayor concentración en la región costa en comparación con la sierra y selva; los cuales están distribuidos en veinte comunidades (Ver Anexo 2). En cuanto a la repetición del grado, se observa un incremento gradual del porcentaje de estudiantes que no aprobaron, pasando de 16.5% en la Ronda 2 a 19.9% en la Ronda 4.

Tabla 8

Distribución de estudiantes según variables de control por rondas

Variable	Ronda 2	Ronda 3	Ronda 4
<i>Región</i>	%		
Costa	50.1	50.4	50.9
Selva	20	20	19.7
Sierra	29.9	29.5	29.4
<i>Aprobó el grado</i>	%		
Sí	83.5	82.3	80.1
No	16.5	17.7	19.9
<i>Asistencia al preescolar</i>	%		
Sí	67.8	67.8	67.8
No	32.2	32.2	32.2
<i>Convive con ambos padres</i>	%		
Sí	66.3	65.1	63.7
No	33.7	34.9	36.3
<i>Tipo de colegio al que asiste</i>	%		
Público	58.2	57.4	55.8
Privado	31.3	32.1	33.5
Híbrido	10.5	10.5	10.7

Adaptado de: Datos de Niños del Milenio, rondas 2-4. Elaboración propia.

La asistencia al preescolar se mantiene constante en todas las rondas, con aproximadamente dos tercios de los estudiantes que sí asistieron a educación inicial. Respecto a la convivencia familiar, se nota una disminución en el porcentaje de estudiantes que viven con ambos padres a lo largo de las rondas. En cuanto al tipo de institución educativa, se observa un ligero descenso en la asistencia a colegios públicos y un aumento en la matrícula de colegios privados, mientras que la modalidad híbrida se mantiene estable.

Variable instrumental: IMC de la madre

En la Tabla 9, se presentan los porcentajes por cada ronda de la variable instrumental (IMC de la madre). Los datos muestran la distribución porcentual del IMC en las categorías definidas en la tabla 5 (desnutrición, peso normal y sobrepeso/obesidad) a lo largo de las tres rondas de evaluación para la

presente investigación. La mayoría de las madres se encuentran en la categoría de sobrepeso/obesidad, con un promedio de 62.2% y sin cambios significativos entre rondas. La proporción de madres con desnutrición se mantiene constante en 20.1%, así como el grupo con peso normal con un promedio de 17.7%.

Tabla 9

Porcentaje por categoría de IMC de la madre por ronda

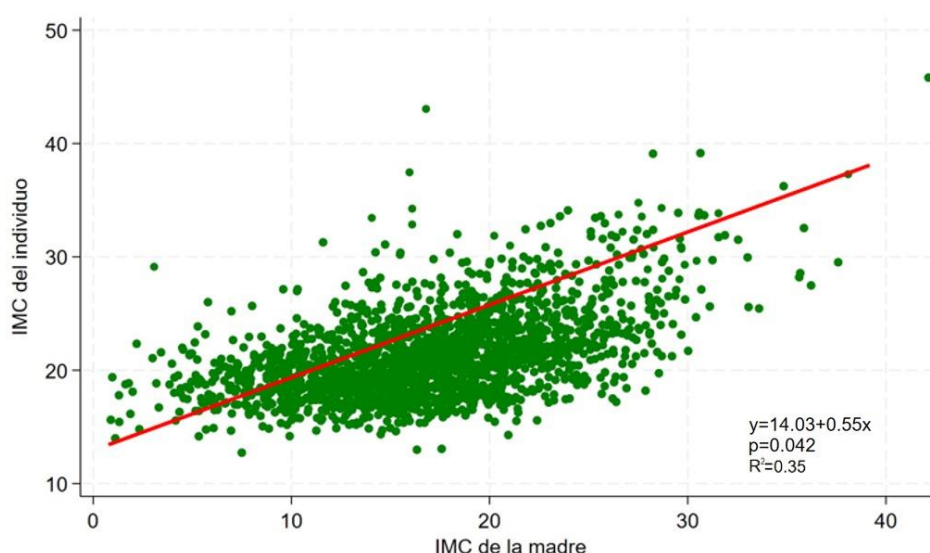
Variable	Ronda 2	Ronda 3	Ronda 4	Total
<i>IMC Madre</i>	%			
Desnutrición	20.1	20.1	20.1	20.1
Normal	17.6	17.6	17.8	17.7
Sobrepeso / <i>Obesidad</i>	62.3	62.3	62.1	62.2

Adaptado de: Datos de Niños del Milenio, rondas 2-4. Elaboración propia.

Adicionalmente, para respaldar la elección del IMC de la madre como variable instrumental, se presenta la relación entre esta variable y el IMC del individuo. La Figura 2 muestra un gráfico de dispersión de estas variables, en el cual se puede observar una tendencia positiva. Esta relación sugiere que, en general, a mayor IMC de la madre, mayor es el IMC del individuo. Esta correlación positiva proporciona evidencia para apoyar la validez del IMC de la madre como variable instrumental.

Figura 2

Gráfico de dispersión del IMC entre madres e hijos



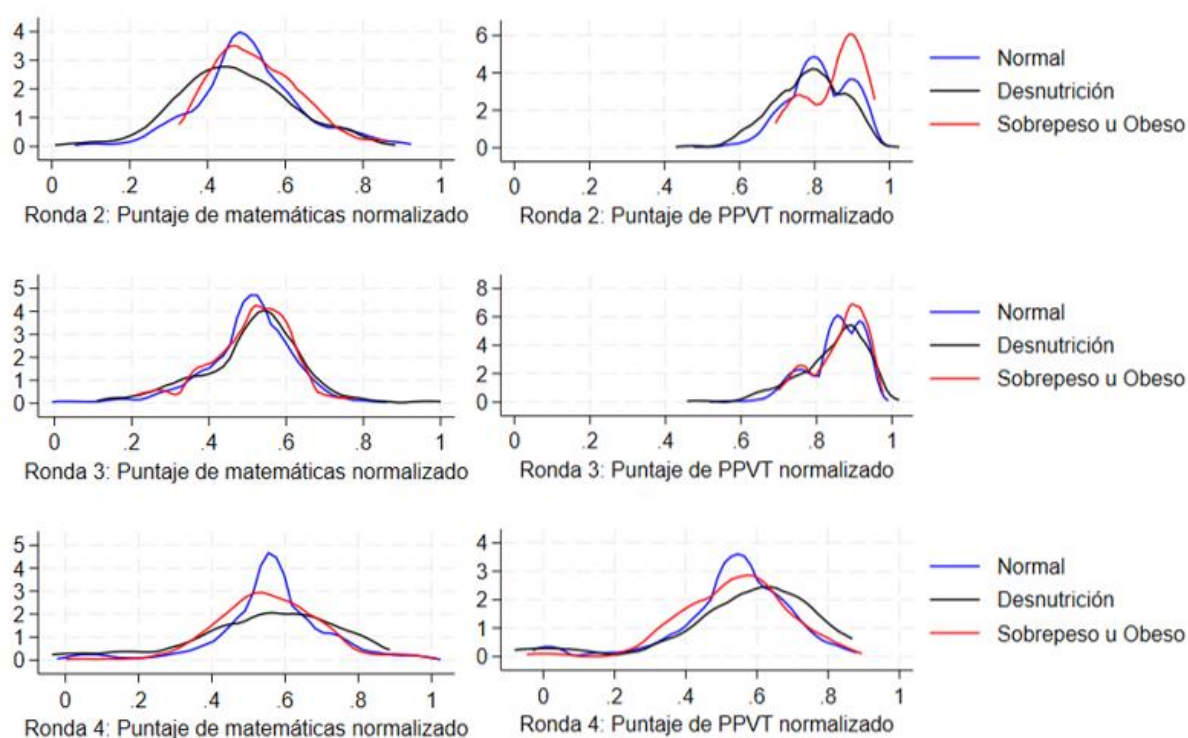
Adaptado de: Datos de Niños del Milenio, rondas 2-4. Elaboración propia.

Sobre la no linealidad del rendimiento y el IMC

Asimismo, se presentan las densidades de Kernel de los puntajes normalizados en matemáticas y PPVT en la Figura 3, diferenciadas según el estado nutricional de los estudiantes (normal, desnutrición y sobrepeso u obesidad) a lo largo de tres rondas³. Las densidades de Kernel son útiles para analizar los datos, ya que permiten visualizar cómo se distribuyen los puntajes dentro de cada grupo, destacando las variaciones y patrones complejos en el rendimiento entre los diferentes estados nutricionales (Wand y Jones, 1995; Scott, 2015).

Figura 3

Densidades de Kernel por rondas de puntajes de matemáticas (Rendimiento matemático) y de PPVT normalizados (Rendimiento comunicativo)



Adaptado de: Datos de Niños del Milenio, rondas 2-4. Elaboración propia.

En la ronda 2, las distribuciones de los puntajes en matemáticas y PPVT muestran diferencias. Para matemáticas, los estudiantes con peso normal presentan un rendimiento superior, evidenciado por un desplazamiento hacia la derecha de su curva de densidad respecto a los grupos con desnutrición y sobrepeso u obesidad. En contraste, en PPVT, el grupo con sobrepeso u obesidad tiene una ventaja,

³ Un gráfico de densidad de kernel no prueba matemáticamente la no linealidad, pero sí brinda indicios fuertes sobre su presencia. Si el gráfico muestra múltiples picos, asimetría o cambios abruptos en la densidad, es probable que una relación no lineal esté presente en los datos.

con su curva desplazada hacia la derecha en comparación con los otros dos grupos. Esto sugiere que el estado nutricional impacta de manera diferente en el rendimiento según el tipo de prueba.

En relación con la ronda 3, las diferencias entre matemáticas y PPVT se reducen. Para matemáticas, el grupo con peso normal sigue mostrando un desempeño superior, mientras que el grupo con desnutrición presenta mayor dispersión en los puntajes, lo que indica variabilidad en el rendimiento. En PPVT, las curvas de los tres grupos están más superpuestas, con una ligera ventaja todavía para el grupo con sobrepeso u obesidad, especialmente en los puntajes altos. Esto sugiere que, a medida que avanza el tiempo, los rendimientos en PPVT tienden a igualarse más entre los grupos en comparación con matemáticas.

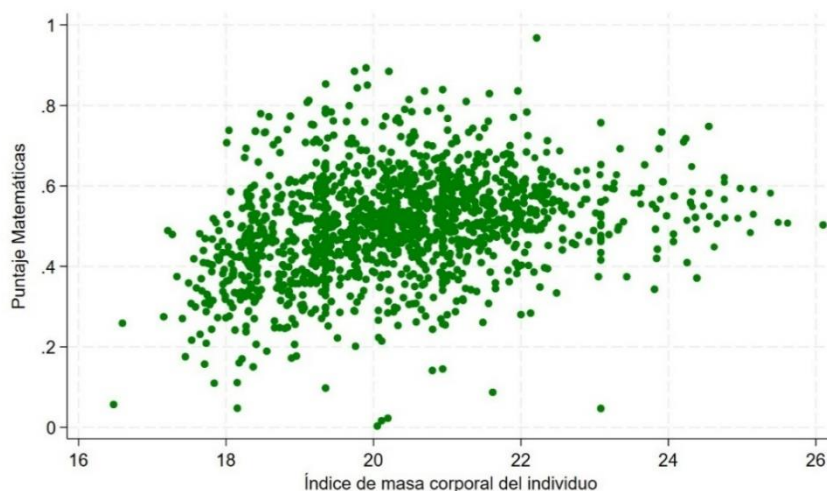
En la ronda 4, se acentúan nuevamente las diferencias. En matemáticas, el grupo con peso normal logra un mejor rendimiento promedio, con su curva desplazada hacia la derecha y más pronunciada, mientras que los grupos con desnutrición y sobrepeso u obesidad muestran densidades más bajas en los puntajes altos. Por otro lado, en PPVT, aunque el grupo con peso normal también muestra un mejor rendimiento, las diferencias entre los grupos son menos marcadas que en matemáticas, lo que indica que la brecha en habilidades verbales es menor que en habilidades matemáticas.

En resumen, las gráficas reflejan un patrón consistente: los estudiantes con peso normal tienen un mejor desempeño promedio en ambas áreas, pero las diferencias son más pronunciadas en matemáticas. En PPVT, los puntajes tienden a ser más similares entre los grupos, particularmente en las rondas intermedias, lo que sugiere que los factores relacionados con el estado nutricional tienen un impacto más evidente en habilidades matemáticas que en habilidades verbales.

Para complementar el análisis y explorar la posible relación no lineal entre el IMC y el rendimiento académico, se realizaron gráficos de dispersión. La Figura 4 muestra la relación entre el IMC y el rendimiento matemático de los estudiantes. El gráfico revela que, para valores bajos de IMC, se observa una leve tendencia positiva en el rendimiento matemático. En el rango de IMC normal, los estudiantes muestran una mayor concentración de puntajes en un nivel intermedio, sin una clara superioridad en los valores más altos. A partir de un IMC superior a 23, la dispersión aumenta y se observa una tendencia negativa pequeña en el rendimiento.

Figura 4

Gráfico de dispersión entre rendimiento matemático e IMC del individuo

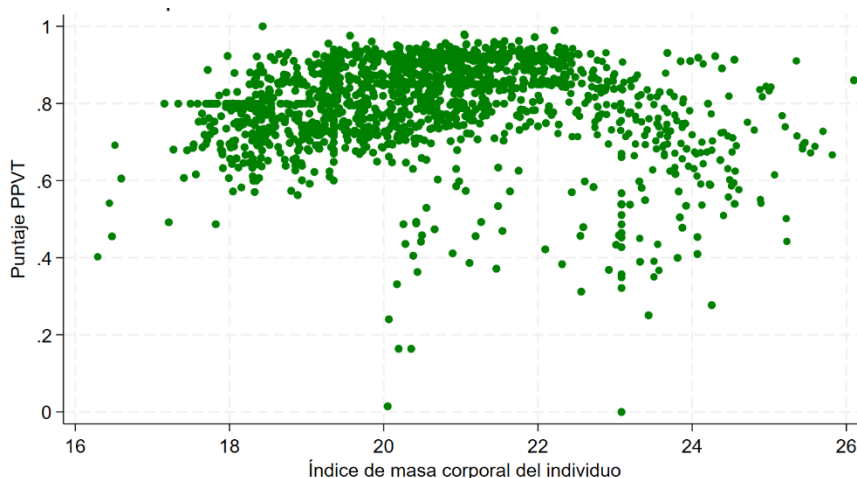


Adaptado de: Datos de Niños del Milenio, rondas 2-4. Elaboración propia.

Por otro lado, la Figura 5 muestra la relación entre el IMC y el rendimiento comunicativo de los estudiantes. El gráfico revela un patrón similar al observado en el rendimiento matemático, pero con algunas diferencias: para valores bajos de IMC (16-18), los puntajes son más dispersos, con algunos valores por debajo de 0.6. A medida que el IMC aumenta hacia el rango de 18 a 22, los puntajes más altos (0.8 - 1.0) se concentran, sugiriendo un mejor desempeño comunicativo en esta zona. Sin embargo, a partir de un IMC superior a 22, la dispersión aumenta y se observan más valores bajos, aunque sin una clara tendencia descendente.

Figura 5

Gráfico de dispersión entre rendimiento comunicativo e IMC del individuo



Adaptado de: Datos de Niños del Milenio, rondas 2-4. Elaboración propia.

Es importante considerar que tanto la Figura 4 como la 5 sugieren una relación no lineal entre el IMC y el rendimiento académico, donde los extremos del IMC (tanto la desnutrición como la obesidad/sobrepeso) están asociados con un menor rendimiento escolar.



Capítulo 3 Marco teórico

Como se ha mencionado anteriormente, en la presente investigación utilizaremos la base de datos de panel de Niños del Milenio desde la ronda 2 (2008) hasta la 4 (2012), en donde se estudian a los individuos desde su niñez hasta la adolescencia⁴. Se propone un modelo teórico basado en el planteamiento de Sabia (2007) para identificar el impacto de la malnutrición, expresado por el IMC, sobre el rendimiento matemático y comunicativo de los individuos. El modelo que propone dicho autor establece una función de utilidad para el individuo con la siguiente forma:

$$U = V(W, I, F, C)$$

Donde U es el rendimiento matemático o comunicativo del individuo, W es el IMC, I son características individuales, F son características familiares y C son características a nivel comunitario. La relación entre W y U dependerá de un umbral máximo y mínimo los cuales definen a las personas que presentan desnutrición y sobrepeso, respectivamente. Esto implica que cuando el IMC supera el umbral máximo de peso normal, es decir, dicha persona tiene sobrepeso u obesidad, la relación será inversamente proporcional, ya que las personas con un IMC elevado a menudo enfrentan problemas de salud física, como enfermedades crónicas, fatiga, problemas de sueño y falta de energía. Tal como se señaló en anteriores capítulos, y según la literatura revisada, estos problemas de salud pueden afectar la capacidad de una persona para concentrarse y aprender en el entorno académico, y por lo consiguiente, el rendimiento académico disminuirá.

Del mismo modo, cuando el IMC es inferior al umbral mínimo de peso normal, es decir, dicha persona se encuentra en un estado de desnutrición, la relación será directamente proporcional, ya que la desnutrición puede llevar a la falta de nutrientes esenciales que son necesarios para el funcionamiento óptimo del cerebro, lo cual genera debilidad, fatiga, problemas cardiacos, anemia, entre otros (Lumun y Kingsley Chukwuma, 2020), lo cual implicaría que al alejarse del umbral de peso normal se esperaría una disminución en el rendimiento. Ambos umbrales generarían una disminución en el rendimiento académico debido a que dificultarían en la concentración y desarrollo del aprendizaje. Por otro lado, las características de la persona a nivel individual están definidas de la siguiente manera:

$$I = o(G, P, O, S, D, K, M, L, Q)$$

Donde G representa si el individuo ha repetido de grado escolar o no, P representa si ha asistido al preescolar o no, O indica si asiste a una escuela pública, privada o híbrida, S indica el sexo de la persona, K representa el nivel máximo educativo alcanzado por el cuidador de la persona, M es

⁴ En la presente investigación, nos referiremos a los últimos años de niñez y adolescencia.

el índice de riqueza normalizado entre 0 y 1, que abarca información de vivienda, servicios y bienes que tiene en el hogar la familia del individuo, L indica si la persona vive con ambos padres y Q indica la región en la que se encuentra el individuo. En base a este modelo teórico, se desarrolla un modelo empírico.



Capítulo 4 Metodología

4.1. Discusión de endogeneidad

La relación entre el IMC y el rendimiento académico presenta desafíos metodológicos debido a la posible violación del supuesto de exogeneidad. Cuando los regresores están correlacionados con el término de error, se generan problemas de identificación que pueden sesgar los resultados y llevar a conclusiones erróneas, sugiriendo efectos causales que en realidad no existen. En este contexto, Sabia (2007)⁵ destaca dos fuentes de endogeneidad en esta relación: la causalidad inversa y factores no observables. Respecto a la causalidad inversa, en la cual no es clara la dirección de la relación causal entre el IMC y el rendimiento académico, plantea los siguientes escenarios:

- i. Un bajo rendimiento académico podría llevar a los estudiantes a comer en exceso como una forma de compensar sus dificultades, aumentando su IMC.
- ii. Un menor rendimiento reduzca el apetito y, por ende, disminuya el IMC.

Asimismo, respecto a los factores no observables (habilidades cognitivas, características psicológicas, como la autoestima, o la predisposición genética) complican la identificación de la verdadera relación entre el IMC y el rendimiento académico. Estas variables, al no poder ser incluidas directamente en el modelo, se incorporan al término de error, generando una correlación espuria con el IMC y, en consecuencia, sesgando las estimaciones.

Para lograr abordar estos desafíos de causalidad, la literatura ha empleado diferentes enfoques metodológicos. Algunos estudios, como los de Neumark-Sztainer (2002) y Krukowski et al. (2009), han empleado controles específicos para los factores psicológicos (como el bullying asociado al sobrepeso), hallazgos que sugieren que el menor rendimiento académico está más relacionado con estos factores sociales, mas no con el IMC. Datar et. al. (2004) y Gable, Chang y Krull (2012), han empleado datos de panel con efectos fijos para examinar la evolución del IMC y el rendimiento académico en distintos grupos de estudiantes. Sus resultados resaltan diferencias de género y apuntan a que habilidades interpersonales y percepciones juegan un rol significativo en el desempeño académico, más allá del peso corporal.

Por otro lado, a partir de la revisión de la literatura, se identifica el uso de variables instrumentales para abordar la problemática de la endogeneidad. Este enfoque se ha desarrollado en dos vertientes: una que utiliza datos relacionados con el IMC como instrumentos y otra que recurre a información genética de los individuos. Con respecto a las instrumentales relacionadas con el IMC, se han empleado variables como el IMC de los padres, la disponibilidad de comida chatarra en el hogar o los patrones de actividad física de los estudiantes. Estas variables son altamente validadas por la

⁵ Basándose en Cawley (2004).

literatura como determinantes relevantes del peso corporal que, sin embargo, no se espera que influyan directamente en el rendimiento académico más allá de su efecto a través del IMC.

En Loken, Mogstad y Wiswall (2012), los autores utilizan el IMC de la madre, así como del padre, como herramientas separadas para determinar si la obesidad tiene una incidencia en el rendimiento académico de los estudiantes de primaria en Noruega, existiendo consistencia entre los resultados. Desde una lógica similar, pero centrados en una muestra de gemelos australianos, se encuentra el estudio de Webbink, Martin y Visscher (2010), quienes emplean el IMC materno y paterno como instrumento, asegurando que dichas variables efectivamente capten la influencia genética del peso sin afectar directamente al rendimiento escolar. Averett y Stifel (2010) analizaron a niños de entre 6 y 13 años en edad escolar, encontrando que aquellos con sobrepeso obtienen un desempeño académico inferior en comparación con sus pares con peso saludable. Este resultado se mantiene consistente al aplicar modelos de efectos fijos individuales y variables instrumentales, utilizando como instrumentos el IMC materno previo al embarazo y su cuadrado.

Finalmente, uno de los estudios más recientes del tema se encuentra en Gwozdz et al. (2015). En este caso también usaron una estrategia instrumental con el IMC parental, pero en una muestra de niños europeos para estudiar el bienestar y rendimiento en los infantes, concluyendo que los resultados son robustos y se pueden generalizar a diferentes contextos culturales y educativos. Asimismo, Cawley et al. (2013) analizaron los patrones de actividad física de los estudiantes, argumentando que las oportunidades limitadas para realizar ejercicio inciden en el peso corporal pero no influyen en el rendimiento académico. Ambos enfoques ofrecen instrumentos que, al estar correlacionados con el IMC, pero no con el rendimiento escolar, son validados empíricamente como adecuados para estudios que buscan identificar relaciones causales.

Por otro lado, las variables sobre información genética han sido exploradas como posibles instrumentos en investigaciones similares como los de Von Hinke et al. (2012) y Tyrrell et al. (2016). Estos autores utilizan diversas variantes polimórficas asociadas al IMC, como aquellas directamente relacionadas con el gen FTO, como instrumentos para estudiar la relación causal entre peso corporal y rendimiento académico. Von Hinke et al. (2012) concluyen que las variantes genéticas que afectan el IMC no influyen directamente en las habilidades cognitivas. Asimismo, Tyrrell et al. (2016), utilizan variante relacionado con la obesidad en el análisis de su impacto en el rendimiento de estudiantes escolares en el Reino Unido, llegando a la conclusión de que los instrumentos utilizados son válidos.

Sin embargo, el uso de estos instrumentos es discutible, ya que existe la preocupación de que los genes que se utilizan como instrumentos se relacionan con otras características cognitivas o de comportamiento que a su vez influyen sobre el rendimiento académico. Un ejemplo de ello es Locke et al. (2015), quienes sugieren que la pleiotropía genética puede invalidar los resultados, ya que los genes que se correlacionan con la obesidad también pueden estar correlacionados con el autocontrol

o la paciencia, características que determinan el rendimiento. Boomsma et al. (2002), alertan acerca del riesgo de instrumentalizar sobre información genética directa en estudios que utilizan variantes polimórficas subrayando la necesidad de verificar que no se correlacionen con otras variables relevantes para el rendimiento escolar.

Por ello, debido a la disponibilidad de datos existentes, se ha decidido utilizar un método de regresión con variable instrumental que permitirá tratar dicha endogeneidad tal como se ha realizado en diversos estudios que evalúan la relación del IMC sobre el rendimiento matemático y comunicativo (por ejemplo: Comuzzie & Allison, 1998; Ding et al., 2009; Norton & Han, 2008). Específicamente, se propone como variables instrumentales el IMC de la madre del individuo, tal como lo abala la literatura previamente mencionada.

4.2. Relación no lineal entre el IMC y el rendimiento académico

Al estimar la relación entre IMC y rendimiento académico, es importante estudiar la no linealidad entre estas variables, la cual se refiere a que el efecto del IMC sobre el rendimiento académico puede no ser constante o proporcional a lo largo de diferentes categorías del IMC. Por ejemplo, pequeños cambios en el IMC podrían tener un efecto negativo en ciertos rangos (como en el caso de desnutrición o sobrepeso/obesidad) y un efecto positivo en el caso de un individuo con peso normal. Con el fin de profundizar en la comprensión de dicha relación, se revisó también literatura que estudia la conexión entre el IMC y los salarios. Esta decisión se tomó debido a que existe una mayor cantidad de estudios en este campo que utilizan modelos no lineales, lo cual puede proporcionar perspectivas valiosas para nuestro análisis. Entre las opciones metodológicas más relevantes se destacan las regresiones cuantílicas, los grados superiores del IMC y los modelos semiparamétricos, cada uno con ventajas específicas según el tipo de relación que se desea explorar.

Primero, las regresiones cuantílicas permiten analizar cómo las diferentes categorías del IMC afectan en el rendimiento académico (matemático o comunicativo), en lugar de centrarse solo en el promedio. Este enfoque es útil para entender si el impacto del IMC es más pronunciado en estudiantes con un rendimiento académico bajo, medio o alto. En lugar de asumir que el efecto del IMC es uniforme, este método examina cómo los efectos varían en los distintos cuartiles de la distribución de los resultados. Por ejemplo, Slade (2017), utilizando un estimador cuantitativo incondicional propuesto por Firpo et al. (2009), encuentra que el efecto del IMC en los salarios varía significativamente según la raza y el género. Para los individuos blancos e hispanos, se observa un efecto de "techo de cristal", donde un mayor IMC se asocia a penalizaciones salariales más fuertes en los niveles salariales altos. En contraste, los hombres negros pueden experimentar una prima salarial con un IMC más alto, mientras que las mujeres negras tienen resultados más uniformes en toda la distribución. Este enfoque podría aplicarse de manera similar para analizar si los estudiantes con peores calificaciones son los más afectados por un alto IMC. Asimismo, Brown (2018) emplea una metodología de regresión

cuantílica para analizar la variabilidad de los efectos de la obesidad sobre los salarios a diferentes niveles de ingresos, permitiendo estimar penalizaciones específicas en cuantiles distintos de la distribución salarial, identificando que las mujeres enfrentan penalizaciones salariales significativamente más altas que los hombres, con una creciente severidad de estos efectos a medida que se avanza en la distribución salarial: las mujeres en el percentil más alto pueden sufrir penalizaciones de hasta el 20%, mientras que los hombres muestran una relación más dispersa y menos consistente.

Aunque existen métodos para combinar regresiones cuantílicas con variables instrumentales, como el enfoque de variables instrumentales de control de función (CFIV), propuesto por Horowitz y Lee (2007), o el estimador de variables instrumentales de regresión cuantílica de Chernozhukov y Hansen (2005), estas no son posibles en este modelo ya que requiere un tamaño de muestra considerable para obtener resultados precisos en cada cuantil (Koenker y Hallock, 2001), lo cual no es posible con nuestra base de datos. Además, este método es sensible a la selección de variables y no aborda problemas de endogeneidad, lo que limita su aplicabilidad en este estudio.

En segundo lugar, el uso de grados superiores del IMC, como términos cuadráticos o cúbicos, es útil para modelar una relación no lineal entre el IMC y el rendimiento académico, es decir, para capturar posibles umbrales críticos en los que los efectos del sobrepeso o la desnutrición cambian radicalmente. Wada y Tekin (2010) emplean el IMC^2 , encuentran una relación positiva para IMC y negativa para IMC^2 , es decir, ganar más peso ayuda a aumentar el puntaje, pero sólo hasta cierto punto, a partir del cual este comienza a disminuir. Debido a que IMC ya es una variable endógena, utilizar grados mayores de este incorporará mayor endogeneidad a la estimación y por lo tanto la necesidad de utilizar nuevas variables instrumentales, las cuales no necesariamente se tienen. Sin embargo, una limitación es el riesgo de multicolinealidad, lo que complica la interpretación de los resultados (Mill-Ferreyra et al., 2018). Además de ello, en el caso específico de este estudio, estos modelos requieren muestras grandes para identificar puntos de inflexión con precisión, lo que no es factible con nuestra base de datos. Asimismo, los gráficos de dispersión de las figuras 4 y 5, presentan una distribución no uniforme de observaciones a lo largo de los diferentes rangos de IMC en relación con el rendimiento, tanto matemático como comunicativo, los cuales tienen una menor densidad en los extremos⁶.

Por último, los modelos semiparamétricos permiten analizar las relaciones entre el IMC y el rendimiento académico sin imponer una forma funcional rígida (Horowitz, 2009). Este enfoque es particularmente relevante para nuestro estudio, ya que como se evidencia en las Figuras 4 y 5, la relación entre el IMC y el rendimiento académico varía de manera no uniforme a lo largo de diferentes

⁶ El análisis preliminar muestra que las observaciones con un IMC bajo u elevado presentan menores valores promedio en rendimiento matemático y PPVT en comparación con los rangos de IMC normal.

rangos de IMC, mostrando patrones distintos para estudiantes con bajo peso, peso normal y sobrepeso/obesidad. Eide et al. (2010) aplicaron modelos semiparamétricos para analizar el impacto de la obesidad en el rendimiento académico, permitiendo capturar heterogeneidades que no se observan con enfoques más rígidos como los modelos paramétricos. Adicionalmente, Gregory y Ruhm (2009) también utilizaron esta metodología y concluyeron que mayores niveles de IMC se relacionan negativamente con menores salarios para las mujeres, a diferencia de los hombres donde la relación es menos clara, ya que depende de los supuestos que se realicen en los modelos. Finalmente, Shimokawa S. (2008) emplea métodos de regresión paramétrica y semiparamétrica para analizar la relación entre el IMC y salarios, hallando que los resultados paramétricos son mixtos y sensibles a la selección de muestras y estrategias de regresión. Sin embargo, los análisis semiparamétricos revelan una penalización salarial tanto para personas con bajo peso como para aquellas con sobrepeso significativo, en ambos géneros, siendo esta penalización más pronunciada en hombres.

Una de las principales ventajas de este enfoque es su flexibilidad, ya que permite detectar patrones complejos y no lineales en los datos sin la necesidad de asumir una relación funcional específica (Robinson, 1987). Sin embargo, estos modelos pueden ser computacionalmente intensivos y más difíciles de interpretar debido a su menor estructura (Horowitz, 2009). No obstante, en la presente investigación emplearemos esta estrategia de análisis debido a que contamos con una base de datos de panel, en la cual se pueden capturar mejor las relaciones no lineales a lo largo del tiempo.

4.3. Modelo semiparamétrico con Kernel

Al elegir la estrategia semiparamétrica, se ha decidido utilizar la regresión de Kernel, pues elimina la suposición de linealidad y modela la expectativa de la variable dependiente como una media ponderada en cada punto de la distribución de la variable independiente (Li y Racine, 2007). En particular, se ha seleccionado el estimador de Kernel de Epanechnikov, uno de los más utilizados en el ámbito del suavizamiento no paramétrico debido a su eficiencia y simplicidad (Epanechnikov, 1969; Gregory y Ruhm, 2009). Comparado con otros estimadores de Kernel, como el de Nadaraya-Watson o el Kernel gaussiano, el de Epanechnikov es óptimo para minimizar el error cuadrático medio integrado (Silverman, 1986; Wasserman, 2006).

En la primera etapa del modelo se utilizará el IMC de la madre como variable instrumental para estimar el IMC del individuo con el fin de abordar el problema de endogeneidad presente en la relación entre el IMC y el rendimiento académico. Luego de ello, en la segunda etapa del modelo, se ajustará, con el estimador de Kernel de Epanechnikov, la función $f(IMC_0^*)$, la cual será una función suavizada que modela la relación no lineal entre el IMC y el rendimiento académico tanto comunicativo como matemático (Horowitz, 2009). La especificación de esta función se encuentra en el Anexo 3. El uso del Kernel de Epanechnikov en el modelo presenta varias ventajas destacando aquella que permite capturar relaciones no lineales entre el IMC y el rendimiento académico sin imponer una forma

funcional rígida, como ocurre en una regresión lineal clásica. Esto ofrece una mayor flexibilidad para adaptar el modelo a la naturaleza de los datos, evitando las limitaciones que surgen al suponer una relación estrictamente lineal.

Además, esta regresión, dada su forma cuadrática truncada, es eficiente para minimizar el sesgo en los extremos de la distribución del IMC. Esta característica es especialmente útil cuando existen pocos datos en los valores más altos o bajos de IMC, ya que permite mejorar la precisión de las estimaciones en esos rangos. Otra ventaja es su menor sensibilidad a los valores atípicos, dado que este tiene soporte limitado, es decir, puede ignorar observaciones extremadamente alejadas de los puntos centrales, lo que reduce el impacto de los outliers y mejora la precisión del modelo.

En conjunto, estas ventajas del estimador semiparamétrico permiten observar cómo el IMC influye en el rendimiento académico, capturando relaciones complejas que un modelo lineal tradicional no detectaría. Por ejemplo, es posible que se encuentre una disminución significativa en el rendimiento académico para individuos con un IMC muy bajo o alto, mientras que aquellos con un IMC "normal" podrían mantener un rendimiento más estable.

4.4. Especificación del modelo

Con el fin de medir el impacto del IMC sobre el rendimiento académico, y aprovechando la naturaleza longitudinal de la base de datos de panel, se empleará, inicialmente un MCO. Sin embargo, teniendo en consideración la problemática de la endogeneidad, se realizará un modelo de efectos fijos a nivel de comunidades⁷ de cada individuo en el modelo de regresión por Mínimos Cuadrados en dos etapas (2SLS) con variable instrumental. Esta metodología permite capturar las diferencias estructurales entre comunidades, asumiendo que sus características no observables permanecen constantes en el tiempo.

Este enfoque ha sido ampliamente aplicado en investigaciones previas. Por ejemplo, Alcázar (2014) incorpora efectos fijos geográficos (por distritos - comunidades) en su modelo de diferencias en diferencias generalizado para controlar por características comunitarias invariables en el tiempo. De forma similar, Pazos et al. (2024) utilizan efectos fijos a nivel de distrito y de fecha de nacimiento para controlar por heterogeneidad no observada tanto temporal como espacial, al analizar el impacto de shocks de lluvia tempranos sobre habilidades cognitivas de niños en Perú.

Asimismo, Arteaga y Glewwe (2014) emplean efectos fijos por comunidad mediante variables dummy para examinar brechas educativas entre niños indígenas y no indígenas, descomponiendo el efecto en componentes asociados al lugar de residencia y a diferencias dentro de la misma comunidad. Otra investigación relevante es la de Hidalgo et al. (2025), que incluye efectos fijos según la región de

⁷ Cabe resaltar que la base de datos de Niños del Milenio define como comunidad a áreas administrativas, que abarcan desde zonas urbanas y barrios hasta centros poblados y caseríos en áreas rurales en donde un conjunto de individuos que comparten una ubicación geográfica específica y pueden estar unidos por intereses, valores, tradiciones y actividades comunes.

nacimiento de la madre, con el objetivo de abordar los efectos regionales persistentes de la violencia política y otras características invariables en el tiempo, definidas a nivel de distrito de nacimiento dentro de cada región. Finalmente, para reforzar la validez de las inferencias estadísticas, se utilizarán errores estándar robustos (clustered) a nivel comunidad, lo que permite corregir posibles problemas de heterocedasticidad y correlación dentro de cada comunidad.

Si bien este enfoque puede resultar en una cierta pérdida de eficiencia en comparación con estimadores que asumen errores independientes, es necesario para obtener inferencias estadísticas válidas dada la estructura de nuestros datos. Considerando el modelo teórico presentado por Sabia (2007), explicada la metodología del modelo de efectos fijos con variable instrumental, definimos nuestro modelo de la siguiente manera:

Primera etapa:

$$IMC_{ict}^* = \beta_0^* + \beta_1^* IMC_{madre\ ict} + \beta_3 X'_{ict} + \gamma_c + E_{ict}$$

Segunda etapa:

$$Rendimiento_{ict} = \beta_0 + \beta_1 (IMC_{ict}^*) + \beta_2 X'_{ict} + \lambda_c + u_{ict}$$

Donde:

- IMC_{ict}^* : es una variable continua que mide el índice de masa corporal del individuo i de la comunidad c en el año t .
- $Rendimiento_{ict}$: es el rendimiento matemático y comunicativo del individuo i de la comunidad c en el año t , normalizado entre 0 y 1 (se estimarán por separado).
- $IMC_{madre\ ict}$: es una variable continua que mide el índice de masa corporal de la madre del individuo i de la comunidad c en el año t .
- X'_{ict} : es el vector de características del individuo i de la comunidad c en el año t , el cual incluirá las siguientes variables:
 - $Grado_{ict}$: es una variable dicotómica que toma el valor de 1 si el individuo i de la comunidad c en el año t no ha repetido el grado cursado actualmente y 0 de lo contrario.
 - $Preescolar_{ic}$: es una variable dicotómica que toma el valor de 1 si el individuo i de la comunidad c ha asistido a un preescolar y 0 de lo contrario.
 - $Colegio_{ict}$: es una variable categórica que toma el valor de 0 si es público, 1 si es privado y 2 si es híbrido (público/privado) para el individuo i de la comunidad c en el año t .
 - $Sexo_{ic}$: es una variable dicotómica que toma el valor de 1 si es mujer y 0 si es hombre para el individuo i en la comunidad c .

- $IndiceRiqueza_{ict}$: es un índice monetario que abarca el acceso a vivienda, servicios y bienes que tiene el individuo i de la comunidad c en el año t , se encuentra normalizado del 0 al 1.
- $Padres_{ict}$: es una variable dicotómica que toma el valor de 1 si individuo i de la comunidad c en el año t convive con ambos padres, de lo contrario toma el valor de 0.
- γ_c : representa el efecto fijo específico en la primera etapa de cada comunidad c a la que pertenece individuo i , controlando todas las características no observadas que son constantes dentro de cada una de ellas, tales como factores geográficos, culturales, socioeconómicos, etc. Este efecto fijo incluirá una matriz de veinte subvariables de comunidades en la regresión, permitiendo capturar el efecto de cada una de ellas.
- λ_c : representa el efecto fijo específico en la segunda etapa de cada región r a la que pertenece individuo i , controlando todas las características no observadas que son constantes dentro de cada una de ellas, tales como factores geográficos, culturales, socioeconómicos, etc. Este efecto fijo incluirá una matriz de veinte subvariables de comunidades en la regresión, permitiendo capturar el efecto de cada una de ellas.
- E_{irt} y u_{irt} son los errores de la primera y segunda etapa, respectivamente.

Asimismo, considerando la no linealidad del IMC, se empleará la función $f(IMC_{irt}^*)$, la cual se encuentra suavizada y que modela la relación no lineal entre el IMC y el rendimiento académico, tanto comunicativo como de matemáticas, mediante la metodología del modelo semiparamétrico con el estimador de Kernel Epachinikov.

Capítulo 5 Resultados

Como se mencionó en la sección metodológica, se estimarán diversas especificaciones para abordar los diferentes problemas relacionados con la medición del impacto del IMC en el rendimiento académico en matemáticas y en comunicación. En primer lugar, se presentarán los resultados del modelo base utilizando mínimos cuadrados ordinarios con efectos fijos. Luego, se mostrarán los resultados del modelo base considerando la variable instrumental del IMC materno y teniendo en cuenta efectos fijos por comunidades. Por último, se expondrán los resultados de la modelación no lineal mediante un modelo semiparamétrico.

5.1. Modelo base: MCO con efectos fijos

En la tabla 10 se presenta el modelo de mínimos cuadrados ordinarios con efectos fijos a nivel de comunidades que busca medir el impacto del IMC en el desempeño académico de los estudiantes considerando los controles a nivel del individuo. Al analizar los resultados se observan tanto similitudes como diferencias en los factores que influyen en el desempeño académico tanto matemático como comunicativo.

Tabla 10

Resultados del modelo de Mínimos Cuadrados Ordinarios con efectos fijos

	(1) Rendimiento Matemático	(2) Rendimiento Comunicativo
IMC	-0.003** (0.001)	-0.004*** (0.001)
Preescolar (1=Asistió)	0.024** (0.008)	0.028*** (0.008)
i.Colegio Público	-0.012 (0.010)	-0.004 (0.009)
i. Colegio Híbrido	0.042 (0.022)	0.049* (0.022)
Sexo (1=Mujer)	-0.009 (0.006)	0.005 (0.006)
Índice de Riqueza	0.196*** (0.018)	0.238*** (0.018)
Grado (1=No ha repetido)	0.015*** (0.002)	0.001 (0.002)
Padres (1=Vive con ambos)	-0.003 (0.007)	0.004 (0.007)
Constante	0.340*** (0.027)	0.728*** (0.026)
EF por comunidad	Si	Si
<i>N</i>	1737	1737
<i>R</i> ²	0.223	0.178

Nota. Errores estándar robustos entre paréntesis. * $p < 0.01$, ** $p < 0.05$, *** $p < 0.1$.

Adaptado de: Datos de Niños del Milenio, rondas 2-4. Elaboración propia.

Los resultados muestran que, respecto al IMC, se tiene un efecto negativo pequeño, pero significativo en ambos tipos de rendimiento, en específico para cada unidad que aumenta el IMC, el rendimiento matemático disminuye en 0.3 puntos porcentuales, mientras que en el rendimiento comunicativo disminuye en 0.4 puntos porcentuales. Asimismo, en relación con las personas que han asistido al preescolar, ambos rendimientos son mayores en promedio para personas que no han tenido preescolar, siendo esta diferencia ligeramente mayor en el área comunicativa.

Por otro lado, los estudiantes provenientes de hogares con mayor índice de riqueza tienden a mostrar, en promedio, un rendimiento más alto tanto en matemáticas como en comunicación. En cuanto a las variables de tipo de institución educativa y género, no se encuentran diferencias significativas en los rendimientos promedio.

5.2. Modelo VI con efectos fijos

Como se explicó en el capítulo de metodología, el modelo de VI tiene como objetivo corregir el sesgo de endogeneidad en la relación entre el IMC y el rendimiento académico, utilizando el IMC de la madre como variable instrumental para el del individuo. Este enfoque permite obtener estimaciones consistentes del impacto del IMC sobre el desempeño académico, mitigando problemas de simultaneidad o de variables omitidas. A continuación, se mostrará la primera etapa del modelo, seguida de la segunda etapa, la cual incorpora el IMC estimado a partir del IMC materno, siendo esta la variable instrumental a emplear.

5.2.1. Primera etapa con VI con efectos fijos

Los resultados de la primera etapa del modelo con variable instrumental se presentan en la Tabla 11, donde el IMC materno se utiliza como variable instrumental para estimar el IMC de los individuos. Como se puede observar, el coeficiente del IMC de la madre es positivo y estadísticamente significativo, lo que sugiere una fuerte relación entre el IMC materno y el de los hijos. El F-estadístico de la primera etapa es 35.163, superando el umbral de 10, lo que indica que el IMC materno es un buen instrumento para el modelo.

Tabla 11*Resultados de la primera etapa del modelo con variable instrumental con efectos fijos*

	(1) IMC
IMC de la madre	1.710** (71.885)
Preescolar (1=Asistió)	0.172 (0.207)
i.Colegio Público	-0.008 (0.027)
i. Colegio Hídrido	0.035 (0.044)
Sexo (1=Mujer)	0.522*** (0.155)
Índice de Riqueza	1.089* (0.455)
Grado (1=No ha repetido)	0.372*** (0.028)
Padres (1=Vive con ambos)	-0.169 (0.173)
Constante	15.246*** (0.648)
EF de comunidad	Si
F	35.163
N	1737
R ²	0.234

Nota. Errores estándar robustos entre paréntesis. * $p < 0.01$, ** $p < 0.05$, *** $p < 0.1$.

Adaptado de: Datos de Niños del Milenio, rondas 2-4. Elaboración propia.

Específicamente, se identifica que un incremento en el IMC de la madre se asocia con un aumento de 1.71 puntos porcentuales en el IMC del individuo. Asimismo, en relación con las variables control, el sexo del estudiante tiende a mostrar, en promedio, un efecto positivo en las mujeres presentando un IMC mayor que los hombres. Asimismo, la riqueza del hogar tiende a mostrar un efecto positivo promedio, al igual que la no repetición de grado. La variable de asistencia al preescolar, el tipo de colegio, y el hecho de vivir con ambos padres no resultan estadísticamente significativas en esta etapa del modelo. La constante es significativa, lo que indica un nivel base considerable del IMC independientemente de las variables incluidas en el modelo.

5.2.2. Segunda etapa con VI con efectos fijos

Los resultados de la segunda etapa del modelo con variable instrumental, presentados en la tabla 12, muestra la relación estimada entre el IMC y el rendimiento matemático (columna 1) y en el rendimiento comunicativo (columna 2).

Tabla 12*Resultados de la segunda etapa del modelo con variable instrumental con efectos fijos*

	(1) Rendimiento Matemático	(2) Rendimiento Comunicativo
IMC	-0.009* (0.015)	-0.008* (0.016)
Preescolar (1=Asistió)	0.022* (0.009)	0.022* (0.009)
i.Colegio Público	-0.021* (0.013)	-0.035*** (0.005)
i. Colegio Híbrido	0.012 (0.011)	0.031 (0.013)
Sexo (1=Mujer)	-0.015 (0.010)	0.001 (0.011)
Índice de Riqueza	0.193*** (0.024)	0.257*** (0.026)
Grado (1=No ha repetido)	0.010 (0.006)	-0.022*** (0.006)
Padres (1=Vive con ambos)	-0.002 (0.007)	0.004 (0.008)
Constante	0.165 (0.230)	0.550* (0.245)
EF de comunidad	Si	Si
<i>N</i>	1737	1737
<i>R</i> ²	0.278	0.226

Nota. Errores estándar robustos entre paréntesis. * $p < 0.01$, ** $p < 0.05$, *** $p < 0.1$. Las variables dependientes (rendimiento matemático y comunicativo) están normalizadas entre 0 y 1.

Adaptado de: Datos de Niños del Milenio, rondas 2-4. Elaboración propia.

En ambas columnas, el IMC muestra un efecto negativo y significativo sobre el rendimiento académico, tanto en el matemático como en el comunicativo. Esto indica que, ante el aumento de una unidad en el IMC, se espera una disminución promedio de 0.9 puntos porcentuales en el rendimiento matemático y de 0.8 puntos porcentuales en el rendimiento comunicativo. En lo que respecta a la asistencia al preescolar se aprecia que, tanto para el rendimiento matemático y comunicativo tienden a ser mayores en promedio para personas que han tenido preescolar. Con respecto al tipo de colegio, se observa que en promedio hay diferencias en el puntaje de matemáticas y comunicación. Para la variable de sexo, en promedio, no se observan diferencias en los puntajes de ninguna de las dos áreas evaluadas. Por otro lado, el índice de riqueza, en promedio, tiene un efecto positivo en ambas áreas, esto indica que los estudiantes de hogares con mayor riqueza obtienen mejores resultados en ambas áreas, especialmente en habilidades lingüísticas. En cuanto a la no repetición del grado, se observa un efecto positivo, en promedio, en el rendimiento matemático; mientras que en el

rendimiento comunicativo es negativo. En promedio, la variable que indica si los estudiantes viven con ambos padres no muestra diferencias en ninguna de las dos áreas evaluadas.

Comparación de resultados MCO y VI

Finalmente, con el fin de evaluar el impacto del IMC en el rendimiento académico de los estudiantes, la Tabla 13 contrasta los resultados obtenidos mediante dos modelos econométricos previamente estimados. En el caso del rendimiento en matemáticas, el coeficiente VI (-0.09 puntos porcentuales) es mayor en magnitud (en valor absoluto) que el coeficiente MCO (-0.03 puntos porcentuales) para el IMC. El hecho de que el coeficiente VI tenga una mayor magnitud indica que el modelo MCO subestima el efecto negativo del IMC sobre el rendimiento matemático.

Por otro lado, en relación con el rendimiento comunicativo, se identifica un patrón similar, específicamente, el coeficiente VI (-0.08 puntos porcentuales), es mayor en magnitud (en valor absoluto) que el coeficiente MCO (-0.04 puntos porcentuales) para el IMC. Del mismo modo que, en el caso del rendimiento matemático, el modelo MCO subestima el efecto negativo del IMC en el rendimiento comunicativo. En otras palabras, la diferencia en los coeficientes indica que el efecto desfavorable del IMC en las habilidades comunicativas del estudiante es más acentuado de lo que el modelo MCO sugirió inicialmente.

Tabla 13

Comparativa entre resultados de los modelos MCO y VI

Variables	Rendimiento Matemáticas		Rendimiento Comunicativo	
	MCO	VI	MCO	VI
IMC	-0.003**	-0.009*	-0.004***	-0.008*
Preescolar (1=Asistió)	0.024**	0.022*	0.028***	0.022*
i. Colegio Público	-0.012	-0.021*	-0.004	-0.035***
i. Colegio Híbrido	0.042	0.012	0.049*	0.031
Sexo (1=Mujer)	-0.009	-0.015	0.005	0.001
Índice de Riqueza	0.196***	0.193***	0.238***	0.257***
Grado (1=No ha repetido)	0.015***	0.010	0.001	-0.022***
Padres (1=Vive con ambos)	-0.003	-0.002	0.004	0.004
Constante	0.340***	0.165	0.728***	0.550*

Nota. Errores estándar robustos entre paréntesis. * $p < 0.01$, ** $p < 0.05$, *** $p < 0.1$. Las variables dependientes (rendimiento matemático y comunicativo) están normalizadas entre 0 y 1.

Adaptado de: Datos de Niños del Milenio, rondas 2-4. Elaboración propia.

Adicionalmente, se realizó el test de Hausman para comparar la eficiencia y consistencia entre los estimadores MCO y VI. Los resultados indican que el estimador VI es más apropiado, lo cual se evidencia en las diferencias significativas entre los coeficientes estimados. Para el rendimiento en matemáticas, el test de Hausman arrojó un estadístico Chi-cuadrado=18.456 ($p = 0.020$), rechazando la hipótesis nula de que los estimadores MCO son consistentes. En el caso del rendimiento

comunicativo, el test de Hausman también rechaza la hipótesis nula con un estadístico Chi-cuadrado=21.07 ($p = 0.001$). Estos resultados respaldan la decisión de utilizar la variable instrumental elegida, ya que las diferencias sistemáticas observadas entre los coeficientes de MCO y VI sugieren la presencia de endogeneidad.

5.2.3. Robustez del instrumento

Para revisar la robustez del instrumento en la segunda etapa de la estimación con variables instrumentales, se realizaron las pruebas correspondientes para evaluar la validez del instrumento utilizado en los modelos. La Tabla 14 presenta los resultados de la prueba de Sargan-Hansen, el cual indica que no se rechaza la hipótesis nula de que los instrumentos utilizados son válidos y no están correlacionados con los errores en el modelo de rendimiento matemático. De manera similar, en el caso del modelo de rendimiento comunicativo, la prueba también sugiere que no hay evidencia de que los instrumentos estén correlacionados con los errores, lo que valida su uso en ambos modelos.

Tabla 14

Test de Sargan-Hansen

Modelo	Chi-cuadrado	Prob > Chi-cuadrado
Rendimiento Matemático	1.24	0.537
Rendimiento Comunicativo	1.29	0.467

Adaptado de: Datos de Niños del Milenio, rondas 2-4. Elaboración propia.

5.3. Modelo semiparamétrico con Kernel

Como se ha mencionado anteriormente, si bien se ha estudiado el efecto del IMC en el rendimiento, pocos estudios han explorado la no linealidad de esta relación. Es por ello que en la tabla 15 se presentan los resultados del modelo semiparamétrico mencionado previamente. Dado que los resultados de la primera etapa, donde se considera la variable instrumental, siguen la misma lógica del modelo base, no se presentan en esta sección.

Tabla 15

Resultados del modelo semiparamétrico del efecto del IMC sobre el rendimiento matemático y comunicativo

	(1) Rendimiento Matemático	(2) Rendimiento Comunicativo
<u>Media</u>		
Rendimiento Matemático	0.498*** (0.004)	
Rendimiento Comunicativo		0.804*** (0.005)
<u>Efecto</u>		
F(IMC*)	-0.018** (0.007)	-0.035*** (0.007)
Índice de Riqueza	0.184*** (0.020)	0.240*** (0.023)
Grado (1=No ha repetido)	0.024*** (0.004)	0.019*** (0.004)
Preescolar (1=Asistió)	0.018 (0.009)	0.018* (0.007)
i.Colegio Público	-0.003 (0.013)	0.010 (0.009)
i. Colegio Híbrido	0.011 (0.025)	0.024 (0.016)
Sexo (1=Mujer)	-0.007 (0.008)	0.002 (0.005)
Padres (1=Vive con ambos)	-0.005 (0.007)	-0.005 (0.006)
EF por comunidad	Si	Si
<i>N</i>	1737	1737
<i>R</i> ²	0.389	0.541

Nota. Errores estándar robustos entre paréntesis. * $p < 0.01$, ** $p < 0.05$, *** $p < 0.1$. Las variables dependientes están normalizadas entre 0 y 1. F(IMC) representa la función no paramétrica del IMC estimada usando el kernel de Epanechnikov.

Adaptado de: Datos de Niños del Milenio, rondas 2-4. Elaboración propia.

Es importante considerar que las medias de los puntajes en matemáticas y en comunicativo son significativas dentro del modelo semiparamétrico, esto indica que los valores centrales de la distribución de estos puntajes no solo reflejan el rendimiento promedio de los estudiantes, sino que también tienen un impacto estadísticamente significativo en el ajuste del modelo. En el contexto, las medias no solo resumen los datos, sino que permiten capturar variaciones clave en la relación entre las variables de interés y las independientes, ajustando tanto los efectos lineales como los no lineales.

El modelo presentado permite diferenciar entre relaciones lineales y no lineales. Por un lado, variables como el índice de riqueza, grado académico y asistencia a preescolar se incluyen de manera

lineal en el modelo, lo que significa que su efecto sobre el rendimiento académico es constante por unidad de cambio. Sin embargo, el efecto del IMC se estima mediante una función no paramétrica, lo que permite capturar posibles relaciones no lineales entre el estado nutricional y el desempeño académico.

Un primer hallazgo es el efecto del IMC en ambas áreas. Se observa que un mayor IMC se asocia con un menor rendimiento académico, con un impacto más pronunciado en las habilidades comunicativas (-3.5 puntos porcentuales) que en matemáticas (-1.8 puntos porcentuales). Este resultado refleja la relación no lineal entre el IMC y el rendimiento, capturada por la función, en contraste con otras variables cuyo efecto es lineal. Asimismo, sobre el índice de riqueza se identifica que los estudiantes provenientes de hogares con mayor riqueza tienden a obtener mejores resultados, efecto que es ligeramente más fuerte en la prueba de habilidades comunicativas. Por otro lado, el grado académico, específicamente el hecho de no haber repetido un año escolar, también presenta un efecto positivo promedio en ambos puntajes. Respecto al acceso a la educación temprana, como la asistencia al preescolar, se identifica que hay un efecto promedio positivo.

5.3.1. Estimación con márgenes predichos

Se realizaron estimaciones de márgenes predichos, los cuales representan los valores promedio esperados del rendimiento académico (matemático y comunicativo) en diferentes niveles específicos del IMC. Estos valores permiten evaluar cómo varía el rendimiento académico en puntos clave de la distribución del IMC: desnutrición (IMC = 16.2, el mínimo valor observado en la muestra), peso normal (IMC = 18.5) y sobrepeso/obesidad (IMC = 25).

En la Tabla 16, se presentan los resultados correspondientes al rendimiento académico en matemáticas. Estos resultados indican que cuando el IMC es 16.2, el rendimiento académico en matemáticas predicho es de 0.504 en la escala normalizada de 0 a 1, lo que representa un nivel intermedio. Para individuos con un IMC = 18.5, el rendimiento académico predicho alcanza su punto más alto de 0.548. En contraste, cuando el IMC es 25, el rendimiento académico predicho cae a su nivel más bajo de 0.494.

Tabla 16

Márgenes del IMC en el rendimiento académico matemático

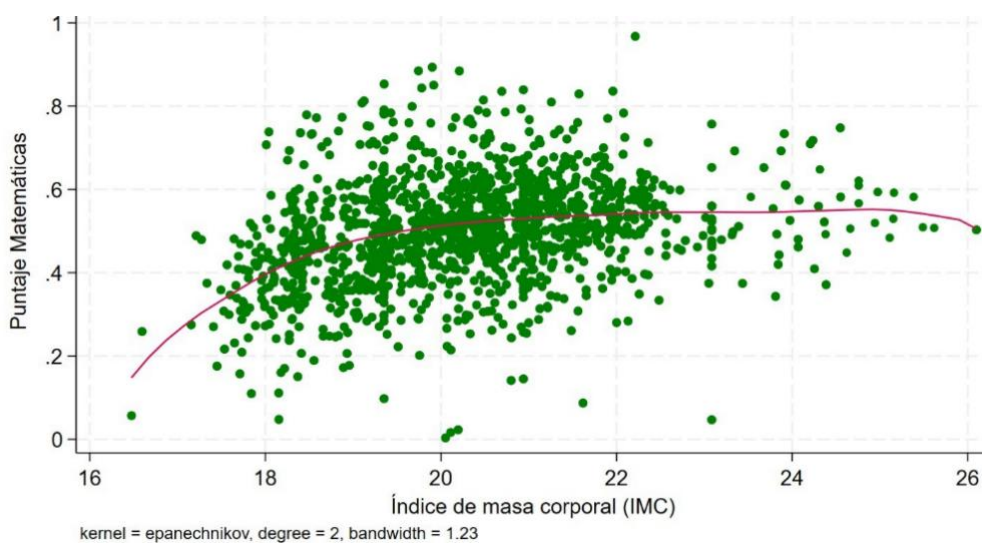
<i>En</i>	<i>Margen observado</i>	<i>Errores robustos</i>	<i>z</i>	<i>P>z</i>
Desnutrición	0.504	0.004	126.690	0.000
Normal	0.548	0.021	26.660	0.000
Sobrepeso/obesidad	0.494	0.005	93.450	0.000

Nota. Desnutrición con valor extremo en IMC =16.2, Normal con valor extremo en 18.5 y sobrepeso/obesidad con valor extremo en 25. Fuente: Datos de Niños del Milenio, rondas 2-4. Adaptado de: Datos de Niños del Milenio, rondas 2-4. Elaboración propia.

Por otro lado, con el fin de poder observar de forma gráfica la relación del IMC y el puntaje en matemáticas, en la Figura 6, se presenta el suavizamiento polinomial de grado dos entre estas variables, el cual complementa de manera visual los resultados obtenidos en la regresión semiparamétrica con Kernel presentada en la tabla 15, específicamente en la primera columna. Este gráfico permite observar cómo varía el rendimiento académico en matemáticas, en función de los distintos niveles de IMC, destacando las diferencias clave a lo largo de la distribución.

Figura 6

Suavizamiento polinomial entre IMC y matemáticas



Adaptado de: Datos de Niños del Milenio, rondas 2-4. Elaboración propia.

La curva muestra una relación no lineal entre el IMC y el puntaje en matemáticas. En los valores más bajos de IMC, correspondientes al rango de desnutrición ($IMC \approx 16.2$), se aprecia que el rendimiento académico es relativamente bajo. A medida que el IMC aumenta hacia el rango normal ($IMC \approx 18.5$), el puntaje en matemáticas también tiende a mejorar, alcanzando su punto máximo alrededor de un IMC de 20. Este aumento sugiere que un IMC más cercano a los valores normales está asociado con un mejor rendimiento académico. Sin embargo, al pasar el punto máximo, la relación comienza a estabilizarse y luego decrece levemente cuando el IMC alcanza el rango superior del peso normal o entra en el rango de sobrepeso ($IMC \geq 25$). Esta disminución en el rendimiento académico es consistente con los márgenes predichos, que indican que los estudiantes con un IMC en estos niveles tienden a tener un puntaje más bajo en matemáticas.

En conjunto, el gráfico ilustra de manera clara que la relación entre el IMC y el rendimiento académico no es lineal. Aunque el rendimiento mejora inicialmente con el incremento del IMC, existe un punto a partir del cual el aumento del IMC deja de ser beneficioso y puede incluso tener un impacto

negativo. Esta representación visual refuerza los hallazgos cuantitativos obtenidos y subraya la importancia de considerar cómo las variaciones en el estado nutricional, especialmente en los extremos de la distribución del IMC, afectan el rendimiento académico.

Por otro lado, la Tabla 17 muestra los márgenes relacionados con el rendimiento académico comunicativo, medido a través del puntaje PPVT. Los resultados indican que, en promedio, las personas con un IMC bajo tienen un rendimiento académico comunicativo predicho de 0.815. En comparación, aquellas con un IMC en el rango normal (IMC = 18.5) presentan un rendimiento académico comunicativo predicho ligeramente mayor (0.823). Finalmente, las personas con un IMC en el límite superior de peso normal o en el rango de sobrepeso (IMC = 25) tienen el rendimiento académico comunicativo más bajo predicho (0.785).

Tabla 17

Márgenes del IMC en el rendimiento académico comunicativo

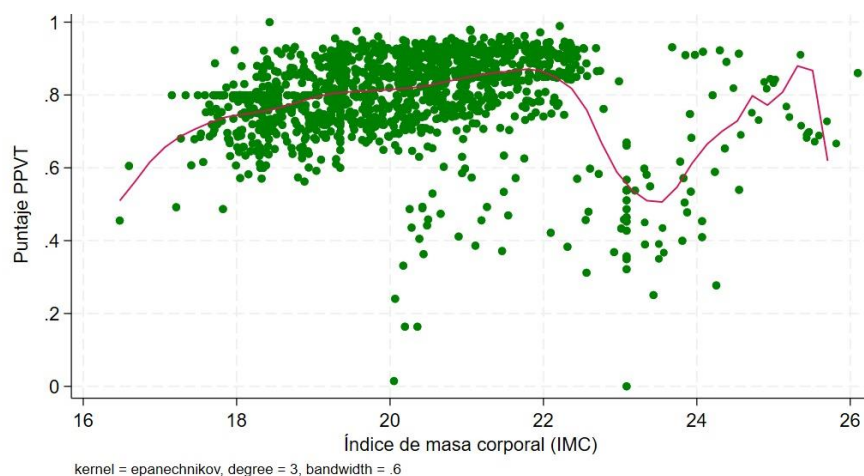
<i>En</i>	Margen observado	Errores robustos	z	P>z
Desnutrición	0.815	0.005	169.440	0.000
Normal	0.823	0.015	56.480	0.000
Sobrepeso/obesidad	0.785	0.008	94.690	0.000

Nota. Desnutrición con valor extremo en IMC =16.2, Normal con valor extremo en 18.5 y sobrepeso/obesidad con valor extremo en 25. Fuente: Datos de Niños del Milenio, rondas 2-4. Adaptado de: Datos de Niños del Milenio, rondas 2-4. Elaboración propia.

Por otro lado, en la Figura 7, se presenta gráficamente el suavizamiento polinomial de grado 3 entre el IMC y el puntaje PPVT proporciona una representación visual de cómo varía el rendimiento académico comunicativo en función del IMC presentada en la tabla 15 en la segunda columna, con el fin de complementar el análisis.

Figura 7

Suavizamiento polinomial entre IMC y PPVT



Adaptado de: Datos de Niños del Milenio, rondas 2-4. Elaboración propia.

A diferencia del puntaje en matemáticas, la relación entre el IMC y el puntaje PPVT presenta una dinámica más compleja, con varias fluctuaciones a lo largo de la distribución. En los niveles más bajos de IMC (indicativos de desnutrición), el rendimiento académico comunicativo es menor, pero incrementa progresivamente conforme el IMC se acerca al rango normal, alcanzando su punto máximo en torno al IMC de 18.5.

A partir de este punto, la curva se estabiliza, mostrando un rendimiento relativamente constante hasta un IMC cercano a 22. Sin embargo, más allá de este nivel, la relación entre el IMC y el puntaje PPVT empieza a mostrar un declive notable, particularmente en los individuos con un IMC superior a 22, lo que coincide con una reducción en el rendimiento comunicativo. Una característica distintiva de este gráfico es la presencia de varios puntos de inflexión hacia el extremo superior de la distribución del IMC, especialmente en los valores correspondientes al sobrepeso y obesidad ($IMC \geq 25$). Esto sugiere una mayor variabilidad en los puntajes comunicativos en estos niveles de IMC, lo que refleja una posible diversidad de factores asociados al rendimiento académico en este rango.

La relación entre el IMC y el rendimiento académico comunicativo, como se refleja en el puntaje PPVT, presenta una tendencia general en la que un IMC en el rango normal está asociado con un rendimiento superior, mientras que los niveles de IMC más altos, así como los más bajos, muestran una tendencia decreciente, confirmando así los márgenes predichos previamente reportados.

5.3.2. Robustez de modelo no lineal

Para asegurar la validez y consistencia de los resultados obtenidos en el modelo de regresión no paramétrica aplicado en esta investigación, se realizaron diversas pruebas de estabilidad y robustez,

las cuales tienen como objetivo evaluar si los efectos estimados permanecen consistentes bajo diferentes especificaciones del modelo y en diferentes subconjuntos de la muestra.

Se llevaron a cabo tres acciones específicas para verificar la robustez del modelo. Una de ellas fue el análisis sobre subgrupos específicos por los rangos del IMC a fin de identificar si los resultados variaban significativamente entre individuos en categorías de desnutrición, peso normal y sobrepeso/obesidad, implementando las variables de control utilizadas en la presente investigación. Finalmente, se realizó un bootstrap con errores robustos considerando un número elevado de repeticiones, lo que permitió obtener estimaciones más precisas a comparación de los errores estándar.

El Anexo 4 presenta los resultados de tres modelos de regresión no paramétrica que evalúan el impacto del IMC en el rendimiento matemático usando el método de mínimos cuadrados en dos etapas (2SLS), donde se emplea el IMC de la madre como variable instrumental. Los resultados están segmentados en tres grupos: desnutrición, peso normal y sobrepeso/obesidad. En los tres grupos, el rendimiento matemático promedio es positivo y estadísticamente significativo. Los estudiantes con un IMC normal presentan el promedio del efecto más alto de rendimiento académico (0.514), seguidos de aquellos que presentan desnutrición (0.467) y, finalmente, los estudiantes con sobrepeso u obesidad (0.427). Esto sugiere que, en promedio, los estudiantes con un IMC normal tienden a obtener mejores resultados en matemáticas, mientras que los estudiantes en la categoría de sobrepeso/obesidad tienden a rendir menos.

El efecto promedio del IMC en el rendimiento matemático varía significativamente según la categoría de peso. En el grupo de desnutrición, el IMC tiene un impacto positivo, lo que sugiere que un aumento de una unidad del IMC dentro de este rango genera un aumento en el rendimiento matemático en 16.8 puntos porcentuales. Para los estudiantes con un IMC normal, el efecto del IMC es pequeño pero positivo (2.4 puntos porcentuales). En contraste, el impacto del IMC en el grupo de sobrepeso/obesidad es negativo (-27.1 puntos porcentuales), indicando que un mayor IMC en este grupo está asociado con una disminución en el rendimiento académico. Estos resultados confirman que la relación entre IMC y rendimiento no es lineal y depende del grupo de IMC al que pertenezca el individuo.

Por otro lado, en el Anexo 5 se presentan los resultados de 3 modelos de regresión que analizan el rendimiento comunicativo de los estudiantes, segmentado en 3 grupos: desnutrición, peso normal y sobrepeso. Se incluyen controles como el índice de riqueza, grado académico, asistencia a preescolar, tipo de colegio, sexo y la convivencia con ambos padres. En los tres grupos, el rendimiento comunicativo promedio es positivo y estadísticamente significativo. Los estudiantes con un IMC normal presentan un promedio más alto en dicho rendimiento (0.811), seguidos por los estudiantes que presentan desnutrición (0.786) y, finalmente, los estudiantes con sobrepeso (0.719). Esto indica

que, en promedio, los estudiantes con un IMC normal tienen un mayor rendimiento en las habilidades comunicativas en comparación con los otros dos grupos.

El efecto del IMC en el rendimiento comunicativo varía según la categoría de peso. En los estudiantes desnutridos, el IMC tiene un impacto positivo (12.5 puntos porcentuales), lo que sugiere que un incremento promedio en el IMC dentro de este rango mejora las habilidades comunicativas. En el grupo de estudiantes con IMC normal, el efecto del IMC es mayor y positivo (32.2 puntos porcentuales). Sin embargo, para los estudiantes con sobrepeso, el impacto del IMC es negativo (-8.3 puntos porcentuales), lo que sugiere que un mayor IMC genera una disminución en el rendimiento comunicativo en este grupo.

En conclusión, los resultados de los modelos para rendimiento matemático y comunicativo demuestran robustez en aspectos clave. En ambos casos, el IMC tiene efectos consistentes y estadísticamente significativos en todas las categorías empleadas. Esto indica que los modelos capturan adecuadamente las relaciones entre las variables de interés y son sólidos frente a diferentes especificaciones. En el caso del rendimiento matemático, el efecto del IMC varía según el grupo: es positivo para los estudiantes con desnutrición y con peso normal, pero negativo para aquellos con sobrepeso, lo que sugiere una no linealidad en la relación entre IMC y desempeño académico. Esta variación confirma que el modelo es robusto y sensible a capturar diferencias importantes en la dinámica entre el estado nutricional y el rendimiento académico. De manera similar, los efectos del IMC en el rendimiento comunicativo son positivos en los grupos con desnutrición y peso normal, pero negativos en el grupo de sobrepeso, reafirmando la coherencia del modelo en diferentes escenarios. Además, las variables de control como el índice de riqueza y grado académico también muestran resultados consistentes entre los modelos de rendimiento matemático y comunicativo.

Conclusiones

De acuerdo con los resultados obtenidos en la presente investigación, se confirma la existencia de una relación no lineal entre el IMC y el rendimiento académico, tanto comunicativo como matemático, en los estudiantes peruanos. Tal como se ha explorado mediante modelos semiparamétricos (Shimokawa, 2008; Horowitz 2009), se observa que los estudiantes con un IMC dentro de los rangos considerados normales tienden a alcanzar un rendimiento académico superior, tanto en matemáticas como en comunicación. En contraste, el impacto negativo en el rendimiento se intensifica en los extremos del IMC, afectando tanto a aquellos estudiantes con desnutrición (IMC por debajo del rango normal), como a quienes presentan sobrepeso u obesidad (IMC por encima del rango normal). Este patrón no lineal es consistente con la evidencia internacional que sugiere efectos diferenciados (Datar et al., 2004; Crosnoe y Muller, 2004; Averett y Stifel, 2010; Gunstad et al. 2008). Con estos resultados, la presente investigación permitió validar la hipótesis planteada, demostrando que la relación entre la malnutrición y el rendimiento de los individuos sigue un patrón no lineal.

Los estudiantes provenientes de familias con un mayor índice de riqueza tienden a mostrar un mejor rendimiento académico, independientemente de su IMC, lo que subraya la importancia del contexto socioeconómico (Filmer y Pritchett, 2001; Vyas y Kumaranayake, 2006). Por otro lado, los resultados no muestran diferencias significativas entre sexos en la relación entre IMC y rendimiento académico. Sin embargo, estudios previos han encontrado tales diferencias. Esta discrepancia podría deberse a diferencias metodológicas o contextuales, y sugiere la necesidad de investigación adicional para entender mejor cómo el impacto del IMC en el rendimiento académico podría variar según el género en diferentes contextos. Sabia (2007) señaló que las adolescentes con sobrepeso tienden a obtener peores resultados académicos que los varones con la misma condición, probablemente debido a una mayor presión social sobre la apariencia física, lo que afecta su bienestar psicológico y, en consecuencia, su desempeño escolar.

Es importante mencionar que en la literatura se encuentra una relación entre el estado nutricional y el rendimiento académico mediada por factores psicosociales, como problemas de autoestima, acoso escolar o exclusión social. Puhl y Latner (2007) encontraron que los adolescentes con sobrepeso son más propensos a ser víctimas de acoso, lo que afecta su autoestima y su motivación para rendir académicamente. Estos efectos pueden explicar por qué en este estudio se observó que los estudiantes con sobrepeso obtuvieron un rendimiento inferior en comparación con aquellos con un IMC normal, a pesar de que las diferencias en otros factores, como el acceso a recursos, fueran mínimas. En este contexto, Neumark-Sztainer (2002) argumentó que el acoso por razones de peso podría tener un impacto más perjudicial en el rendimiento académico que el propio sobrepeso, destacando la importancia de los factores sociales en la relación entre el IMC y el rendimiento escolar.

Recomendaciones

Para futuras investigaciones, se recomienda explorar instrumentos alternativos o complementarios que presenten mayor variación temporal, como la cantidad de actividad física que realizan los individuos o los hábitos alimenticios de las familias, que podrían proporcionar una perspectiva más dinámica de esta relación. Asimismo, sería útil desarrollar métodos que aprovechen mejor la naturaleza longitudinal de los datos, como modelos de efectos mixtos o técnicas de análisis de series temporales, que podrían capturar de manera más precisa la evolución temporal de dicha relación. Adicionalmente, se recomienda que en estudios futuros se analicen las relaciones vinculadas a las diferencias entre las zonas rurales y urbanas, ya que los estudiantes de las regiones de sierra y selva presentan mayores desafíos más elevados. A su vez, investigar las razones detrás de las disparidades en la accesibilidad de los servicios de salud y educación conllevan a la malnutrición y como está importancia influye en el rendimiento escolar. De igual manera, estudiar los efectos a largo plazo de la malnutrición no abordada en el desarrollo cognitivo utilizando datos longitudinales más recientes.

Por otro lado, durante el desarrollo del estudio se identificaron algunas limitaciones metodológicas, en específico, la ausencia de efectos fijos a nivel individual y la potencial endogeneidad de algunas variables de control. Esto sugiere que futuras investigaciones podrían beneficiarse de estas especificaciones u otras alternativas que permitan controlar mejor las características no observables. Del mismo modo, es importante considerar que los hallazgos de este estudio son relevantes para la formulación de políticas públicas. Por ejemplo, se podrían implementar programas integrales de nutrición escolar a nivel nacional que atiendan a la desnutrición y al sobrepeso/obesidad, ya que ambos extremos de la malnutrición influyen negativamente en el rendimiento académico. Estos deben ser priorizados en las zonas rurales, donde las tasas de malnutrición suelen ser más altas y los recursos son limitados. De esta manera, se podría garantizar una intervención eficaz con la finalidad de mejorar el rendimiento escolar en las comunidades más vulnerables.

Asimismo, para complementar dichos programas es necesario incluir la educación nutricional en el currículo escolar, tanto para estudiantes como para sus padres, con el objetivo de promover hábitos alimenticios saludables desde la infancia. Paralelamente, se deben desarrollar campañas de concientización sobre el impacto de la malnutrición en el rendimiento académico y la salud a largo plazo, fomentando una cultura de alimentación adecuada. Finalmente, resulta clave la implementación de programas de monitoreo del IMC en niños y adolescentes, permitiendo identificar casos de malnutrición y brindar intervenciones tempranas, garantizando así un seguimiento continuo del estado nutricional de los estudiantes y su impacto en el aprendizaje.

Referencias

- Alcázar, L. (2014). Impactos del programa Juntos sobre el empoderamiento de la mujer. (Avances de Investigación, 19). Lima: GRADE Group for the Analysis of Development. <https://nbn-resolving.org/urn:nbn:de:0168-ssoar-56528-2>
- Arteaga, I., & Glewwe, P. (2014). Achievement gap between indigenous and non-indigenous children in Peru: An analysis of Young Lives survey data (Young Lives Working Paper No. 130). Young Lives, Oxford Department of International Development, University of Oxford. https://www.younglives.org.uk/sites/default/files/migrated/YLWP130_Arteaga_Achievement%20Gaps%20Peru.pdf
- Averett, S. L., & Stifel, D. C. (2010). Race and gender differences in the cognitive effects of childhood overweight. *Applied Economics Letters*, 17(17), 1673-1679. DOI: 10.1080/13504850903251256
- Benavides, M. (2002). Para explicar las diferencias en el rendimiento en matemática de cuarto grado en el Perú urbano: Análisis de resultados a partir de un modelo básico. En J. Rodríguez & S. Vargas (Eds.), *Análisis de resultados y métodos de las pruebas CRECER 1998* (pp. 93-107). Ministerio de Educación.
- Boomsma, D., Busjahn, A., & Peltonen, L. (2002). Classical twin studies and beyond. *Nature reviews. Genetics*, 3(11), 872-882. <https://doi.org/10.1038/nrg932>
- Brown, C., & Routon, P. W. (2018). On the distributional and evolutionary nature of the obesity wage penalty. *Economics and Human Biology*, 28, 160-172. <https://doi.org/10.1016/j.ehb.2017.10.001>
- Cawley, J. (2004). The impact of obesity on wages. *Journal of Human Resources*, 39(2), 451-474. <https://doi.org/10.2307/3559022>
- Cawley, J., Frisvold, D., & Meyerhoefer, C. (2013). The impact of physical education on obesity among elementary school children. *Journal of health economics*, 32(4), 743-755. <https://doi.org/10.1016/j.jhealeco.2013.04.006>
- Chang, S. M., Walker, S. P., Grantham-McGregor, S., & Powell, C. A. (2010). Early childhood stunting and later fine motor abilities. *Developmental medicine and child neurology*, 52(9), 831-836. <https://doi.org/10.1111/j.1469-8749.2010.03640.x>
- Chetty, R., Friedman, J. N., Hilger, N., Saez, E., Whitmore Schanzenbach, D., & Yagan, D. (2011). How does your kindergarten classroom affect your earnings? Evidence from Project STAR. *The Quarterly Journal of Economics*, 126(4), 1593-1660. <https://doi.org/10.1093/qje/qjr041>
- Comuzzie, A. G., & Allison, D. B. (1998). The search for human obesity genes. *Science*, 280(5368), 1374-1377. <https://doi.org/10.1126/science.280.5368.1374>

- Crookston, B. T., Schott, W., Cueto, S., Dearden, K. A., Engle, P., Georgiadis, A., Lundeen, E. A., Penny, M. E., Stein, A. D., & Behrman, J. R. (2013). Postinfancy growth, schooling, and cognitive achievement: Young Lives. *The American journal of clinical nutrition*, 98(6), 1555–1563. <https://doi.org/10.3945/ajcn.113.067561>
- Crosnoe, R., & Muller, C. (2004). Body mass index, adolescent achievement, and school context: Examining the educational experiences of adolescents at risk of obesity. *Journal of Health and Social Behavior*, 45(4), 393-407. <https://doi.org/10.1177/002214650404500403>
- Cueto, S. (2005). Height, weight and educational achievement in rural Peru. *Food and Nutrition Bulletin*, 26(2), S251-S260. <https://doi.org/10.1177/15648265050262S216>
- Cueto, S., Guerrero, G., León, J., Seguín, E. & Muñoz, I. (2009). Explaining and overcoming marginalization in education: a focus on ethnic/language minorities in Perú. *Global Monitoring Report 2010*. Recuperado de: <http://unesdoc.unesco.org/images/0018/001865/186589e.pdf>
- Currie, J., & Stabile, M. (2006). Child mental health and human capital accumulation: The case of ADHD. *Journal of Health Economics*, 25(6), 1094-1118. <https://doi.org/10.1016/j.jhealeco.2006.03.001>
- Chernozhukov, V., & Hansen, C. (2005). An IV model of quantile treatment effects. *Econometrica*, 73(1), 245-261. <https://doi.org/10.1111/j.1468-0262.2005.00570.x>
- Chinyoka, K. (2014). Impact of poor nutrition on the academic performance of grade seven learners: A case of Zimbabwe. *International Journal of Learning & Development*, 4(3), 73-84. <https://doi.org/10.5296/ijld.v4i3.6169>
- Datar, A., Sturm, R., & Magnabosco, J. L. (2004). Childhood overweight and academic performance: National study of kindergartners and first-graders. *Obesity Research*, 12(1), 58-68. <https://doi.org/10.1038/oby.2004.9>
- Ding, W., Lehrer, S. F., Rosenquist, J. N., & Audrain-McGovern, J. (2009). The impact of poor health on academic performance: New evidence using genetic markers. *Journal of Health Economics*, 28(3), 578-597. <https://doi.org/10.1016/j.jhealeco.2008.11.006>
- Eide, E. R., Showalter, M. H., & Goldhaber, D. D. (2010). The relation between children's health and academic achievement. *Children and Youth Services Review*, 32(2), 231-238. <https://doi.org/10.1016/j.childyouth.2009.08.019>
- Epanechnikov, V. A. (1969). Non-parametric estimation of a multivariate probability density. *Theory of Probability & Its Applications*, 14(1), 153-158. <https://doi.org/10.1137/1114019>
- Filmer, D., & Pritchett, L. H. (2001). Estimating wealth effects without expenditure data—or tears: An application to educational enrollments in states of India. *Demography*, 38(1), 115-132. <https://doi.org/10.1353/dem.2001.0003>

- Fondo de las Naciones Unidas para la Infancia. (2022). Estado mundial de la infancia 2021: En mi mente: Promover, proteger y cuidar la salud mental de la infancia. UNICEF.
- Gable, S., Krull, J. L., & Chang, Y. (2012). Boys' and girls' weight status and math performance from kindergarten entry through fifth grade: a mediated analysis. *Child development*, 83(5), 1822–1839. <https://doi.org/10.1111/j.1467-8624.2012.01803.x>
- Gregory, C. A., & Ruhm, C. J. (2011). Where does the wage penalty bite? NBER Chapters, in: *Economic Aspects of Obesity*, pages 315-347, National Bureau of Economic Research, Inc. <https://doi.org/10.3386/w14984>
- Gunnarsdottir, C. N., & Nines, M. K. (2015). The relationship between body mass index and academic achievement. *Economic Inquiry*, 53(2), 1233-1247. <https://doi.org/10.1111/ecin.12176>
- Gunstad, J., Spitznagel, M. B., Paul, R. H., Cohen, R. A., Kohn, M., Luyster, F. S., Clark, R., Williams, L. M., & Gordon, E. (2008). Body mass index and neuropsychological function in healthy children and adolescents. *Appetite*, 50(2-3), 246-251. <https://doi.org/10.1016/j.appet.2007.07.008>
- Gwozdz, W., Sousa-Poza, A., Reisch, L. A., Bammann, K., Eiben, G., Kourides, Y., Kovács, É., Lauria, F., Konstabel, K., Santaliestra-Pasias, A. M., Vyncke, K., & Pigeot, I. (2015). Peer effects on obesity in a sample of European children. *Economics and human biology*, 18, 139–152. <https://doi.org/10.1016/j.ehb.2015.05.002>
- Hanushek, E. A., & Woessmann, L. (2008). The role of cognitive skills in economic development. *Journal of Economic Literature*, 46(3), 607-668. <https://doi.org/10.1257/jel.46.3.607>
- Hidalgo, A., Rojas, P., & Torero, M. (2025). The long shadow of conflict on human capital: Intergenerational transmission of mothers' exposure to civil war in Peru. *Journal of Development Economics*, 160, 102987. <https://doi.org/10.1016/j.jdeveco.2025.103468>
- Hioui, M., Ahami, A., Aboussaleh, Y., & Rusinek, S. (2016). The relationship between nutritional status and educational achievements in the rural school children of Morocco. *Journal of Neurology and Neurological Disorders*, 3(1), 1-4. <https://doi.org/10.15744/2454-4981.3.101>
- Hoddinott, J., Alderman, H., Behrman, J. R., Haddad, L., & Horton, S. (2013). The economic rationale for investing in stunting reduction. *Maternal & child nutrition*, 9 Suppl 2 (Suppl 2), 69–82. <https://doi.org/10.1111/mcn.12080>
- Horowitz, J. L., & Lee, S. (2007). Nonparametric instrumental variables estimation of a quantile regression model. *Econometrica*, 75(4), 1191-1208. <https://doi.org/10.1111/j.1468-0262.2007.00786.x>
- Horowitz, J. L. (2009). Semiparametric and nonparametric methods in econometrics. *Annual Review of Economics*, 1(1), 225-244. <https://doi.org/10.1007/978-0-387-92870-8>
- Instituto Nacional de Estadística e Informática. (2020). Perú: Encuesta Demográfica y de Salud Familiar- ENDES 2019. INEI.

- Kar, B. R., Rao, S. L., & Chandramouli, B. A. (2008). Cognitive development in children with chronic protein energy malnutrition. *Behavioral and brain functions : BBF*, 4, 31. <https://doi.org/10.1186/1744-9081-4-31>
- Kaestner, Robert & Grossman, Michael. (2008). Effects of Weight on Children's Educational Achievement. *Economics of Education Review*. 28. 651-661. <https://doi.org/10.1016/j.econedurev.2009.03.002>
- Koenker, R., & Hallock, K. F. (2001). Quantile regression. *Journal of Economic Perspectives*, 15(4), 143-156. <https://doi.org/10.1257/jep.15.4.143>
- Krukowski, R. A., West, D. S., Philyaw Perez, A., Bursac, Z., Phillips, M. M., & Raczynski, J. M. (2009). Overweight children, weight-based teasing and academic performance. *International journal of pediatric obesity : IJPO : an official journal of the International Association for the Study of Obesity*, 4(4), 274–280. <https://doi.org/10.3109/17477160902846203>
- Li, Q., & Racine, J. S. (2007). *Nonparametric econometrics: Theory and practice*. Princeton University Press.
- Lobstein, T., & Brinsden, H. (2019). *Atlas of childhood obesity*. World Obesity Federation.
- Locke, A. E., Kahali, B., Berndt, S. I., Justice, A. E., Pers, T. H., Day, F. R., & Hirschhorn, J. N. (2015). Genetic studies of body mass index yield new insights for obesity biology. *Nature*, 518(7538), 197-206. <https://doi.org/10.1038/nature14177>
- Loken, K., Mogstad, M., & Wiswall, M. (2012). What Linear Estimators Miss: The Effects of Family Income on Child Outcomes. *American Economic Journal: Applied Economics*, 4 (2): 1–35. DOI: 10.1257/app.4.2.1
- Lumun, N., & Kingsley Chukwuma, N. (2020). Malnutrition, human growth and development in Nigeria: A need for community engagement and social protection interventions. *International Journal of Social Sciences and Humanities Reviews*, 10(1), 171-178.
- Maureira Cid, F., Palma Gajardo, E., Medina Saavedra, R., Segueida Lorca, Á., Valenzuela Contreras, L., & Flores Ferro, E. (2019). Incidencia de la antropometría, práctica de actividad física, estilos de aprendizaje, motivos, actitudes y estrategias de aprendizaje sobre el rendimiento académico de estudiantes de Santiago de Chile. *Retos*, 36, 497-502. <https://doi.org/10.47197/retos.v36i36.69895>
- Mendoza, J. L. (2019). Efectos de la desnutrición en el rendimiento escolar de los niños en el Perú. *Pensamiento Crítico*, 24(1), 79-102. <https://doi.org/10.15381/pc.v24i1.16562>
- Mill-Ferreira, E., Cameno-Carrillo, V., Saúl-Gordo, H., & Camí-Lavado, M. (2018). Estimación del índice de masa corporal con base en la circunferencia braquial, para pacientes con discapacidad permanente o transitoria. *Medicina de Familia SEMERGEN*, 44(5), 304-309. <https://doi.org/10.1016/j.semerg.2017.08.002>

- Monge, A., Campana, Y., & Gutiérrez, L. G. (2017). El efecto de la desnutrición crónica infantil sobre el desempeño educativo de los estudiantes peruanos de segundo de primaria [Informe Final]. Consorcio de Investigación Económica y Social.
- Must, A., & Strauss, R. S. (1999). Risks and consequences of childhood and adolescent obesity. *International Journal of Obesity and Related Metabolic Disorders*, 23(S2), S2-S11. <https://doi.org/10.1038/sj.ijo.0800852>
- Neumark-Sztainer, D., Falkner, N., Story, M., Perry, C., Hannan, P. J., & Mulert, S. (2002). Weight-teasing among adolescents: correlations with weight status and disordered eating behaviors. *International journal of obesity and related metabolic disorders : journal of the International Association for the Study of Obesity*, 26(1), 123–131. <https://doi.org/10.1038/sj.ijo.0801853>
- Norton, E. C., & Han, E. (2008). Genetic information, obesity, and labor market outcomes. *Health Economics*, 17(9), 1089-1104. <https://doi.org/10.1002/hec.1386>
- Pazos, N., Favara, M., Sánchez, A., Scott, D., & Behrman, J. (2024). Long-term effects of early life rainfall shocks on foundational cognitive skills: Evidence from Peru. *Economics and Human Biology*, 54, 101407. <https://doi.org/10.1016/j.ehb.2024.101407>
- Puhl, R. M., & Latner, J. D. (2007). Stigma, obesity, and the health of the nation's children. *Psychological Bulletin*, 133(4), 557-580. <https://doi.org/10.1037/0033-2909.133.4.557>
- Ramsey, D., & Muskin, P. R. (2013). Vitamin deficiencies and mental health: How are they linked? Identifying and correcting deficiencies can improve brain metabolism and psychopathology. *Current Psychiatry*, 12(1), 37-44.
- Richardson, A. J., & Ross, M. A. (2000). Fatty acid metabolism in neurodevelopmental disorder: A new perspective on associations between ADHD, dyslexia, dyspraxia and the autistic spectrum. *Prostaglandins, Leukotrienes and Essential Fatty Acids*, 63(1-2), 1-9. <https://doi.org/10.1054/plaf.2000.0184>
- Robinson, P. M. (1988). Semiparametric Econometrics: A Survey. *Journal of Applied Econometrics*, 3(1), 35–51. <http://www.jstor.org/stable/2096563>
- Rojas, N. F., Li, C. A., Dávila, S. K., & Alva, M. R. (2015). El estado nutricional y su impacto en los logros de aprendizaje. *Ciencia Amazónica*, 5(2), 115-120. <https://doi.org/10.22386/ca.v5i2.97>
- Sabia, J. J. (2007). The effect of body weight on adolescent academic performance. *Southern Economic Journal*, 73(4), 871-900. <https://doi.org/10.1002/j.2325-8012.2007.tb00809.x>
- Sabia, J. J., & Rees, D. I. (2015). Body weight, mental health capital, and academic achievement. *Review of Economics of the Household*, 13(4), 653-684. <https://doi.org/10.1007/s11150-014-9272-7>

- Saintila, J., & Rodríguez, M. (2016). Estado nutricional y rendimiento académico en escolares de 7 a 14 años de la Institución Educativa Mi Jesús, Lurigancho, Lima. *Revista Científica de Ciencias de la Salud*, 9(2), 63-71. <https://doi.org/10.17162/rccs.v9i2.656>
- Sánchez-Cruzat, D., Llopis-Morales, A., & Morales-Suárez-Varela, M. (2017). Obesity and academic performance in primary school students: A multicenter study. *International Journal of Environmental Research and Public Health*, 14(5), 501. <https://doi.org/10.3390/ijerph14050501>
- Scott, D. W. (2015). *Multivariate density estimation: Theory, practice, and visualization* (2nd ed.). John Wiley & Sons. DOI:10.1002/9781118575574
- Selamawit, W. K., Amene, A. K., & Samson, K. D. (2021). Adolescents' nutritional status and its association with academic performance in South Ethiopia: A facility-based cross-sectional study. *BMC Nutrition*, 7(1), 52. <https://doi.org/10.1186/s40795-021-00420-8>
- Shimokawa S. (2008). The labour market impact of body weight in China: a semiparametric analysis, *Applied Economics*, 40:8, 949-968, DOI: 10.1080/00036840600771239
- Silverman, B. W. (1986). *Density estimation for statistics and data analysis*. Chapman & Hall/CRC.
- Slade, P. (2017). Body mass and wages: New evidence from quantile estimation. *Economics and Human Biology*, 27, 223-240. <https://doi.org/10.1016/j.ehb.2017.07.001>
- Themane, M. J., Monyeki, K. D., Nthangeni, M. E., Kemper, H. C. G., & Twisk, J. W. R. (2003). The relationship between health (malnutrition) and educational achievements (Maths and English) in the rural children of South Africa. *International Journal of Educational Development*, 23(6), 637-643. [https://doi.org/10.1016/S0738-0593\(03\)00063-4](https://doi.org/10.1016/S0738-0593(03)00063-4)
- Tyrrell, J., Jones, S. E., Beaumont, R., Astley, C. M., Lovell, R., Yaghootkar, H., & Frayling, T. M. (2016). Genetic evidence for causal relationships between obesity and depression. *Nature Human Behaviour*, 1(2), 1-9. <https://doi.org/10.1038/s41562-016-0024>
- Vashti-Dominguez, A. H., & Halili, B. L. B. (2018). Food for thought: The socioeconomic impact of child malnutrition and maternal health on the academic performance of Filipino school children. *European Journal of Sustainable Development*, 7(4), 361-371. <https://doi.org/10.14207/ejsd.2018.v7n4p361>
- Victora, C. G., Adair, L., Fall, C., Hallal, P. C., Martorell, R., Richter, L., Sachdev, H. S., & Maternal and Child Undernutrition Study Group (2008). Maternal and child undernutrition: consequences for adult health and human capital. *Lancet* (London, England), 371(9609), 340-357. [https://doi.org/10.1016/S0140-6736\(07\)61692-4](https://doi.org/10.1016/S0140-6736(07)61692-4)
- Von Hinke Kessler Scholder, S., Davey Smith, G., Lawlor, D. A., Propper, C., & Windmeijer, F. (2012). The effect of fat mass on educational attainment: examining the sensitivity to different identification strategies. *Economics and human biology*, 10(4), 405-418.

- <https://doi.org/10.1016/j.ehb.2012.04.015>
- Von Hinke Kessler Scholder, S., Davey Smith, G., Lawlor, D. A., Propper, C., & Windmeijer, F. (2016). Genetic markers as instrumental variables. *Journal of Health Economics*, 45, 131-148. <https://doi.org/10.1016/j.jhealeco.2015.10.007>
- Vyas, S., & Kumaranayake, L. (2006). Constructing socio-economic status indices: How to use principal components analysis. *Health Policy and Planning*, 21(6), 459-468. <https://doi.org/10.1093/heapol/czl029>
- Wada, R., & Tekin, E. (2010). Body composition and wages. *Economics & Human Biology*, 8(2), 242-254. <https://doi.org/10.1016/j.ehb.2010.02.001>
- Wasserman, N. (2006). Stewards, agents, and the founder discount: Executive compensation in new ventures. *Academy of Management Journal*, 49(5), 960-976. <https://doi.org/10.5465/amj.2006.22798177>
- Wand, M.P., & Jones, M.C. (1994). *Kernel Smoothing* (1st ed.). Chapman and Hall/CRC. <https://doi.org/10.1201/b14876>
- Webbink, D., Martin, N. G., & Visscher, P. M. (2010). Does education reduce the probability of being overweight? *Journal of Health Economics*, 29(1), 29-38. <https://doi.org/10.1016/j.jhealeco.2009.11.013>
- Wisniewski, S. L. (2017). Childhood obesity among the poor in Peru: Are there implications for cognitive outcomes? *Economics and Human Biology*, 26, 190-203. <https://doi.org/10.1016/j.ehb.2017.02.003>
- Zaini, M. Z., Lim, C. T., Low, W. Y., & Harun, F. (2005). Effects of nutritional status on academic performance of Malaysian primary school children. *Asia Pacific Journal of Public Health*, 17(2), 81-87. <https://doi.org/10.1177/101053950501700204>

Anexos

Anexo 1

Registro de observaciones eliminadas en la base de datos

	<i>Missings</i>	<i>Outliers</i>	<i>Repetidos</i>
Ronda 2	99	21	15
Ronda 3	75	17	7
Ronda 4	32	15	9
Total	206	53	31

Adaptado de: Niños del Milenio, ronda 2-4. Elaboración propia.

Anexo 2

Comunidades de Niños de Milenio

Identificación	Departamento	Nombre anonimizado	Breve descripción
1	Tumbes	Tumbes	Una ciudad pequeña en la costa norte, en la región de Tumbes.
2	Piura	Sullana	Un área rural costera pobre en la región de Piura.
3	Piura	Morropón	Un área rural muy pobre en la sierra norte, en la región de Piura.
4	Amazonas	Chachapoyas	Un área rural muy pobre en el norte de la región de Amazonas.
5	San Martín	Rioja	Un área rural pobre en la región de San Martín.
6	San Martín	San Martín	Una ciudad de tamaño mediano en la región de San Martín.
7	Cajamarca	Cajamarca	Una ciudad de tamaño mediano en la sierra norte, en la región de Cajamarca.
8	La Libertad	Trujillo	Un asentamiento humano en las cercanías de una ciudad de tamaño mediano.
9	Ancash	Huaylas	Un área rural muy pobre en la sierra central.
10	Ancash	Huaraz	Una ciudad de tamaño mediano en la sierra central, en la región de Ancash.
11	Huánuco	Dos de Mayo	Un área rural muy pobre en el centro de la sierra, en la región de Huánuco.
12	Lima	San Juan de Lurigancho	Un distrito urbano grande situado en el norte de la ciudad capital, Lima.
13	Lima	Ate	Un distrito urbano grande situado en el este de la ciudad capital, Lima.
14	Lima	Villa María del Triunfo	Un distrito urbano grande situado en el sur de la ciudad capital, Lima.
15	Junín	Satipo	Un área rural pobre en la parte amazónica de la región de Junín.
16	Ayacucho	Huamanga	Una comunidad rural pobre en la sierra sur central.
17	Ayacucho	Lucanas	Una comunidad rural pobre en la sierra sur central.

Identificación	Departamento	Nombre anonimizado	Breve descripción
18	Apurímac	Andahuaylas	Una comunidad rural pobre en la sierra sur.
19	Arequipa	Camaná	Una pequeña ciudad en la costa sur, región de Arequipa.
20	Puno	Juliaca	Una ciudad de tamaño mediano en la sierra sur, en la región de Puno.

Adaptado de: Diseño y métodos del estudio Niños del Milenio en el Perú.

Anexo 3

Especificación de la función Kernel de Epanechnikov

$$f(IMC_0^*) = \frac{\sum_{i=1}^n K\left(\frac{IMC_{irt}^* - IMC_0^*}{h}\right) Rendimiento_{irt}}{\sum_{i=1}^n K\left(\frac{IMC_{irt}^* - IMC_0^*}{h}\right)}$$

Donde:

- $K\left(\frac{IMC_{irt}^* - IMC_0^*}{h}\right)$: Es el Kernel de Epanechnikov, que toma la forma $K(x) = \frac{3}{4}(1 - x^2)$ para $|x| \leq 1$ y $K(x) = 0$ para $|x| > 1$.
- h : Es el ancho de banda, que controla el grado de suavizamiento. Valores más pequeños de h dan menos suavizamiento (más precisión local), y valores mayores suavizan más (reduciendo el ruido).
- IMC_{irt}^* : es una variable continua que mide el índice de masa corporal del individuo i de la región r en el año t estimada en la primera etapa del modelo, donde el índice de masa corporal de la madre se utiliza como variable instrumental para predecir el IMC del individuo.
- IMC_0^* : es una variable puntual de IMC^* en donde se estimará la función f .
- $Rendimiento_{irt}$: es el rendimiento matemático y comunicativo del individuo i de la región r en el año t , normalizado entre 0 y 1 (se estimarán por separado).

Anexo 4*Resultados de análisis de subgrupos para el rendimiento matemático*

	(1) Rendimiento Matemático Desnutrición	(2) Rendimiento Matemático Normal	(3) Rendimiento Matemático Sobrepeso/Obeso
<u>Media</u>			
<i>Rendimiento matemático</i>	0.467*** (0.004)	0.514*** (0.005)	0.427*** (0.004)
<u>Efecto</u>			
<i>F(IMC*)</i>	0.168* (0.002)	0.024** (0.007)	-0.271** (0.007)
<i>Índice de Riqueza</i>	0.231*** (0.026)	0.197*** (0.018)	0.178*** (0.025)
<i>Grado (1=No ha repetido)</i>	0.041*** (0.005)	0.025*** (0.003)	0.021*** (0.004)
<i>Preescolar (1=Asistió)</i>	0.023 (0.012)	0.020* (0.008)	0.022* (0.010)
<i>i. Colegio Público</i>	-0.001 (0.018)	0.002 (0.006)	0.009 (0.010)
<i>i. Colegio Híbrido</i>	-0.005 (0.037)	0.007 (0.011)	0.006 (0.020)
<i>Sexo (1=Mujer)</i>	-0.008 (0.011)	-0.001 (0.012)	0.004 (0.004)
<i>Padres (1=Vive con ambos)</i>	-0.008 (0.011)	-0.004 (0.009)	-0.003 (0.011)
<i>EF por comunidad</i>	Si	Si	Si
<i>N</i>	221	930	586
<i>R²</i>	0.315	0.209	0.293

Nota. Errores estándar robustos entre paréntesis. * $p < 0.01$, ** $p < 0.05$, *** $p < 0.1$. Las variables dependientes (rendimiento matemático y comunicativo) están normalizadas entre 0 y 1.

Adaptado de: Datos de Niños del Milenio, rondas 2-4. Elaboración propia.

Anexo 5*Resultados de análisis de subgrupos para el rendimiento comunicativo*

	(1)	(2)	(3)
	Rendimiento Comunicativo Desnutrición	Rendimiento Comunicativo Normal	Rendimiento Comunicativo Sobrepeso
Media			
<i>Rendimiento comunicativo</i>	0.786*** (0.003)	0.811*** (0.003)	0.719*** (0.003)
Efecto			
<i>F(IMC*)</i>	0.125* (0.002)	0.322*** (0.005)	-0.083*** (0.014)
<i>Índice de Riqueza</i>	0.153*** (0.015)	0.237*** (0.015)	0.278*** (0.032)
<i>Grado (1=No ha repetido)</i>	0.023*** (0.002)	0.019*** (0.003)	0.019** (0.007)
<i>Preescolar (1=Asistió)</i>	0.012 (0.008)	0.025** (0.006)	0.039*** (0.009)
<i>i. Colegio Público</i>	-0.008 (0.011)	0.015 (0.008)	0.035*** (0.008)
<i>i. Colegio Híbrido</i>	-0.023 (0.021)	0.033* (0.016)	0.075*** (0.016)
<i>Sexo (1=Mujer)</i>	-0.001 (0.006)	0.011* (0.006)	0.031*** (0.008)
<i>Padres (1=Vive con ambos)</i>	-0.016*** (0.004)	0.005 (0.007)	0.008 (0.009)
EF por comunidad	Si	Si	Si
<i>N</i>	221	930	586
<i>R²</i>	0.317	0.359	0.273

Nota. Errores estándar robustos entre paréntesis. * $p < 0.01$, ** $p < 0.05$, *** $p < 0.1$. Las variables dependientes (rendimiento matemático y comunicativo) están normalizadas entre 0 y 1.

Adaptado de: Datos de Niños del Milenio, rondas 2-4. Elaboración propia.