

¿Por qué han mejorado los rendimientos educativos en el Perú?: Un contraste de hipótesis para el periodo 2013 – 2015

(resultados preliminares)

Alvaro Monge Zegarra
Yohnny Campana Morales
Lais Grey Gutierrez

Categoría: PM

Ámbito: Nacional

Área de Estudio: A1 – Gestión de Recursos Naturales y Crecimiento Económico

Tema: T4 - Inversión Pública, Canon y Mejoramiento de la Calidad Educativa

- 1. Introducción**
- 2. Literatura previa**
- 3. Inferencias preliminares**
- 4. Los datos**
- 5. Metodología**
- 6. Resultados**
- 7. Hallazgos preliminares**

Motivación. Entre los años 2012 y 2013 se observa un cambio de tendencia en los rendimientos educativos, ¿este cambio se debe exclusivamente a factores de oferta y demanda contemporáneos o también influyó la mejora en nutrición ocurrida 4 años antes?

Hipótesis. La mejora en la nutrición experimentada en el pasado ha influido de manera positiva e importante en el cambio de tendencia en los indicadores de rendimiento académico observados hoy por lo que la sostenibilidad del progreso educativo depende críticamente de los progresos adicionales en esta variable del desarrollo infantil temprano.

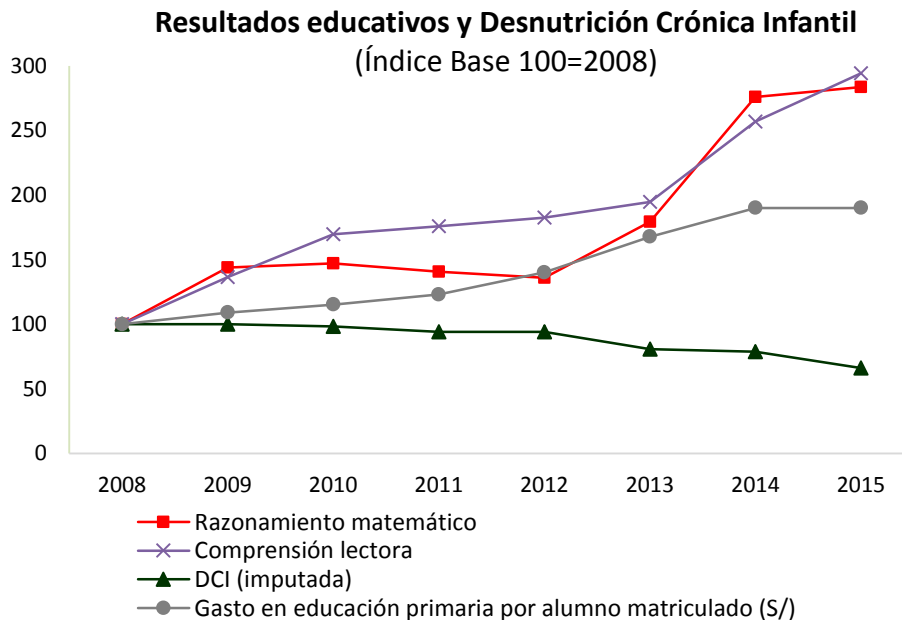
Objetivo general. Generar evidencia sobre los factores que han incidido en la mejora de los indicadores de desempeño académico en los últimos años, considerando aspectos asociados al stock nutricional de los niños y aislando la influencia de otros factores potencialmente explicativos, como la oferta escolar o factores de demanda contemporáneos. Asimismo, en la medida de lo posible, indagar el grado de responsabilidad de cada uno de estos aspectos en el cambio de tendencia experimentado.

Objetivos específicos:

- Medir el efecto individual que tiene la desnutrición infantil sobre el rendimiento académico de los estudiantes.
- Indagar si es que el efecto marginal de las políticas a nivel de oferta educativa es mayor si es que ocurre en un contexto de mejor stock nutricional.
- Realizar ejercicios que permitan identificar heterogeneidades críticas: género, grupo étnico, ámbito geográfico, entre otras.
- Sobre la base de los resultados anteriores extraer lecciones de política sobre los espacios de articulación entre los sectores educativo y salud.

- Coleman y otros (1966): diferencias en la oferta educativa no generan diferencias en resultados académicos. Sin embargo, Hanushek y Kain (1972), Hanushek (1996 y 2003), Greenwald y otros (1996), Ravkin y otros (2005) critican informe Coleman y argumentan que el contexto educativo sí tiene influencia (directa e indirecta) en el proceso de aprendizaje del estudiante
- Cueto (2004) y Cueto y otros (2010) comentan que regularidad empírica muestra dos tipos de determinantes: oferta (características del servicio educativo) y demanda (características de los estudiantes y sus familias); con diferente nivel de influencia dependiendo del contexto analizado y del tipo de rendimiento (razonamiento matemático o verbal). Al respecto, Banco Mundial (1999), Brunner y Elacqua (2003), Carrasco (2007) y Beltrán y Seinfeld (2011) con diferente énfasis proponen la idea que los factores de demanda son mas importantes en países desarrollados y los factores de oferta en países en desarrollo.
- Sin embargo, evidencia en Perú (Caro, 2003; Cueto, 2007; Benavides y Rodriguez, 2006) muestra que la influencia de la oferta educativa en rendimientos es débil explicando menos del 10% de la variabilidad de los resultados en pruebas estandarizadas.
- Nuevas explicaciones: Outes y otros (2011) y Dercon y Sanchez (2011) muestran evidencia sobre influencia de la nutrición en habilidades cognitivas (test Peabody) y no cognitivas (autoestima, aspiración, etc.); y Sanchez (2013) muestra relación entre habilidades no cognitivas en habilidades cognitivas. Estos son nuestros antecedentes más cercanos, pero proponemos mejoras como (i) vincular directamente las variables de política (desnutrición en el periodo infantil temprano y rendimiento educativo mediante pruebas estandarizadas de segundo de primaria); (ii) usar desnutrición pasada (en línea con modelo teórico).

Desnutrición y rendimiento: evolución reciente

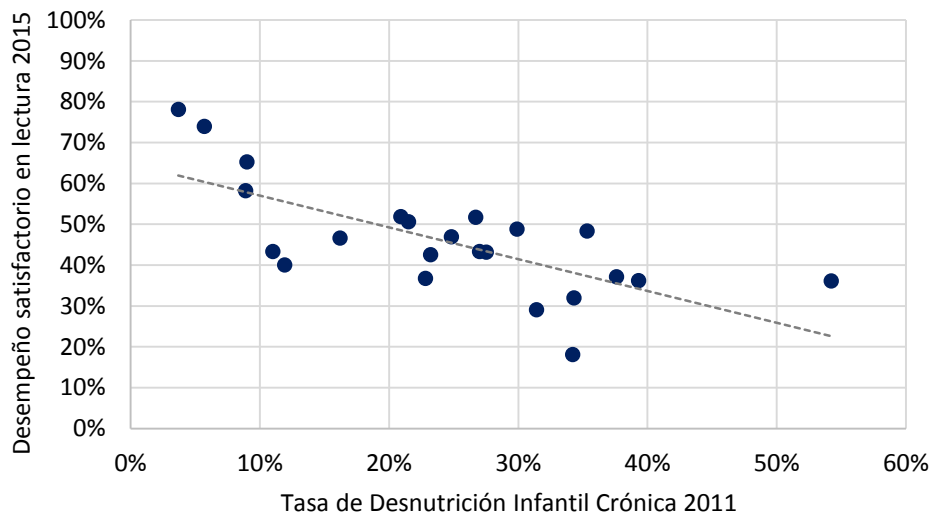


- En el Perú, la mitad de los niños no comprenden lo que leen y 3 de cada 10 no pueden resolver problemas matemáticos básicos. Lo anterior determina que de manera recurrente el estudiante peruano se ubique en el tramo inferior de la prueba PISA.
- Sin embargo, hay tendencias positivas que vale la pena destacar, sobre todo a partir de los años 2012 (matemáticas) y 2013 (lenguaje). En los últimos 3 años la comprensión lectora ha aumentado casi 20 puntos y el razonamiento matemático casi 15 puntos.
- Existen diferentes explicaciones para este cambio de tendencia
 - La pobreza se reduce en el periodo
 - Se observan mejoras en variables de oferta (aproximadas por el gasto en educación primaria per-cápita)
 - El stock nutricional de los niños entrantes a partir de 2012 – 2013 era mejor.

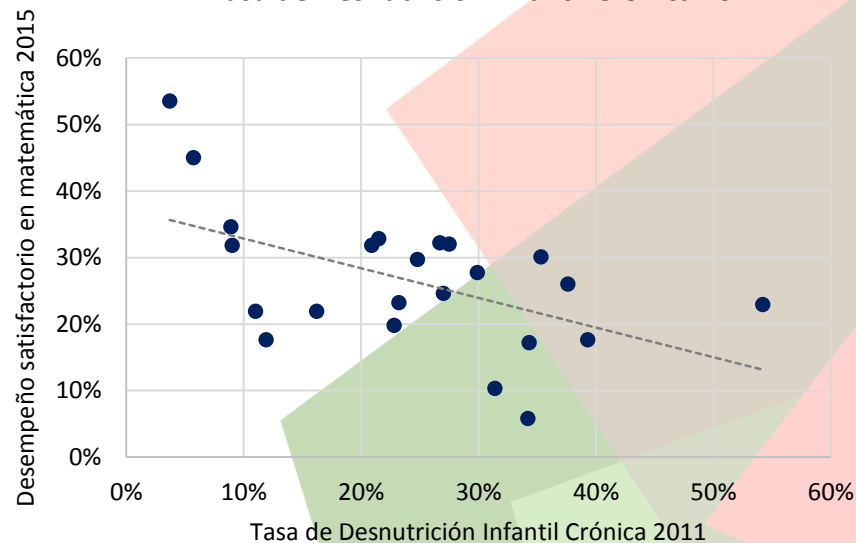
Determinantes previos: Formación de capital humano

- **Parece existir una relación negativa entre el desempeño educativo de los niños y su stock nutricional:** las regiones que presentan las menores tasas de DCI (2011) son aquellas donde los alumnos muestran un mejor desempeño en comprensión lectora y matemáticas (2015).

**Desempeño satisfactorio en lectura en 2015 vs.
Tasa de Desnutrición Infantil Crónica 2011, según regiones**



**Desempeño satisfactorio en matemáticas 2015 vs.
Tasa de Desnutrición Infantil Crónica 2011**

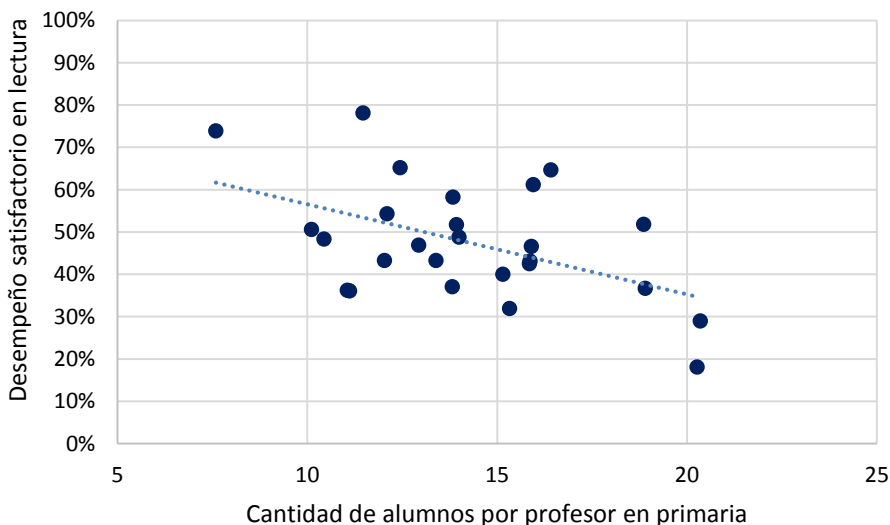


Fuente: ENDES 2011 y ECE 2015. Elaboración propia.

Determinantes de oferta: Cantidad de alumnos por profesor

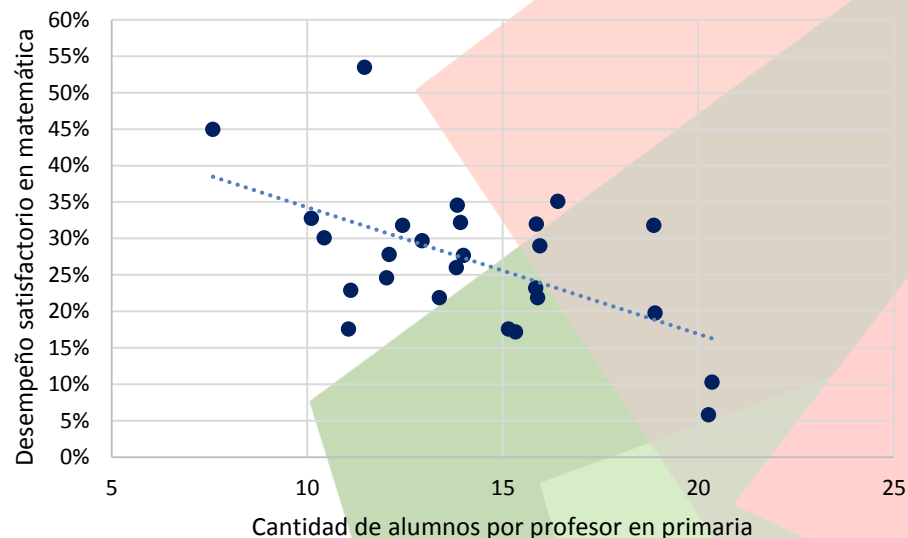
- **Parece existir una relación negativa entre el desempeño educativo de los niños y la cantidad de alumnos por profesor en primaria:** las regiones que cuentan con una menor cantidad de profesores por alumno en primaria son aquellas donde los alumnos muestran un mejor desempeño en comprensión lectora y matemáticas (2015).

**Desempeño satisfactorio en lectura vs.
Cantidad de alumnos por profesor, según regiones 2015**



Fuente: ECE y Censo Escolar 2015. Elaboración propia.

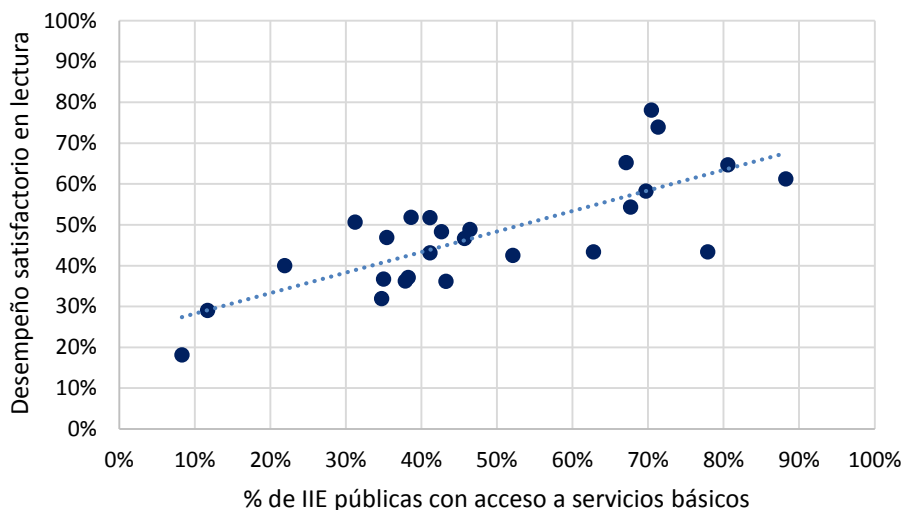
**Desempeño satisfactorio en matemática vs.
Cantidad de alumnos por profesor, según regiones 2015**



Determinantes de oferta: Acceso a servicios básicos

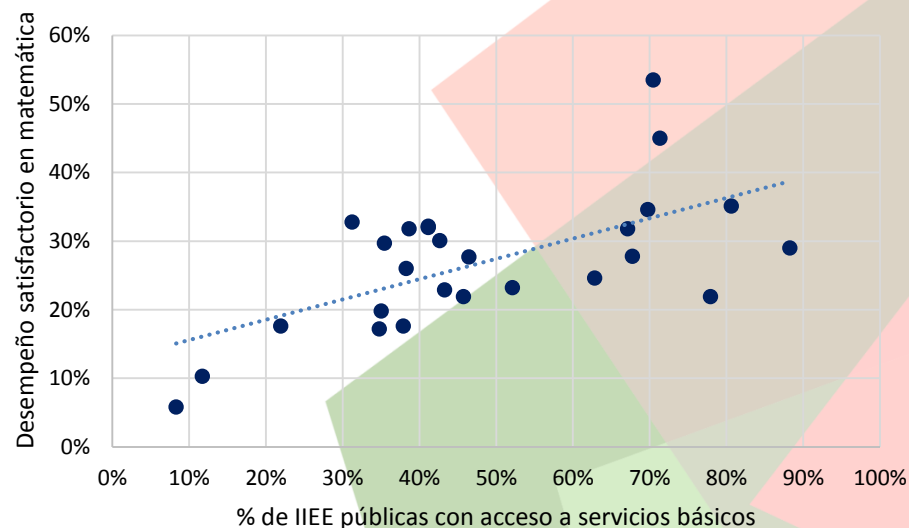
- **Parece existir una relación positiva entre el desempeño educativo de los niños y el acceso a servicios básicos de las escuelas:** las regiones que cuentan con una mayor proporción de IIEE con los tres servicios básicos (agua, saneamiento y electricidad) son aquellas donde los alumnos muestran un mejor desempeño en comprensión lectora y matemáticas (2015).

**Desempeño satisfactorio en lectura vs.
Acceso a servicios básicos, según regiones 2015**



Fuente: ECE y Censo Escolar 2015. Elaboración propia.

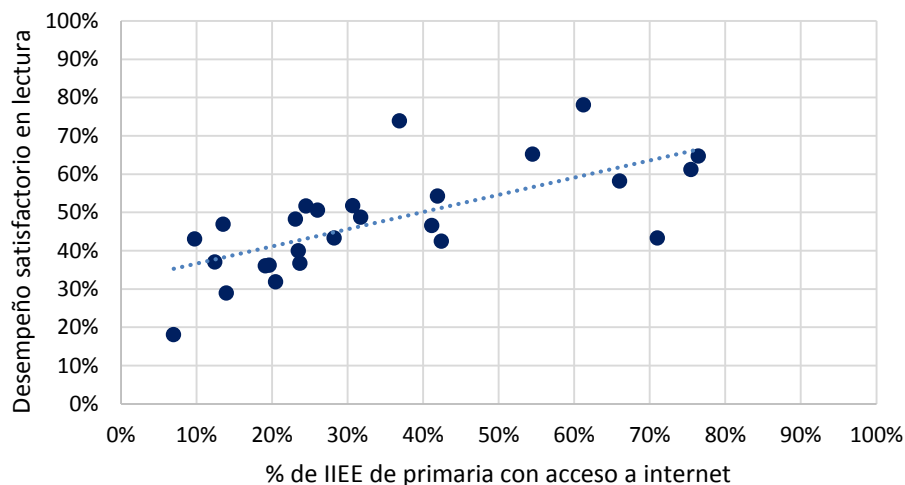
**Desempeño satisfactorio en matemática vs.
Acceso a servicios básicos, según regiones 2015**



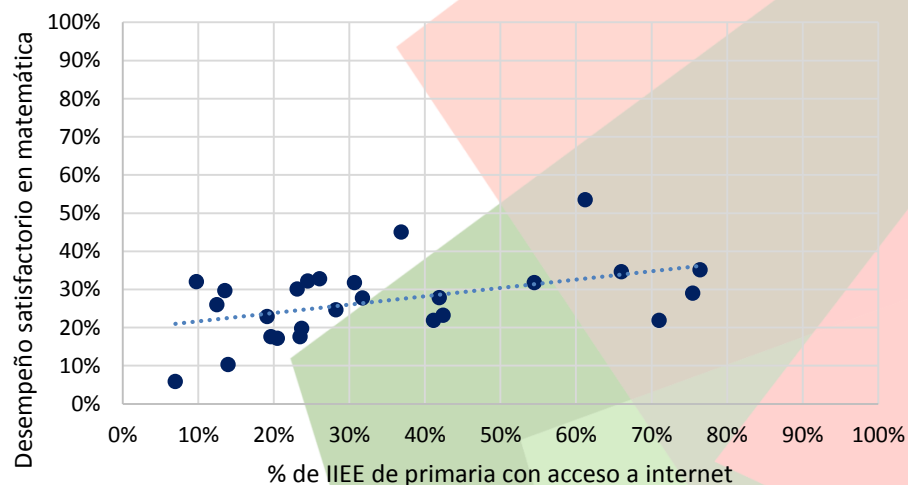
Determinantes de oferta: Acceso a internet

- **Parece existir una relación positiva entre el desempeño educativo de los niños y el acceso internet en las escuelas:** las regiones que cuentan con una mayor proporción de IIEE de primaria con acceso a internet son aquellas donde los alumnos muestran un mejor desempeño en comprensión lectora y matemáticas (2015).

Desempeño satisfactorio en lectura vs. Acceso a internet, según regiones 2015



Desempeño satisfactorio en matemática vs. Acceso a internet, según regiones 2015

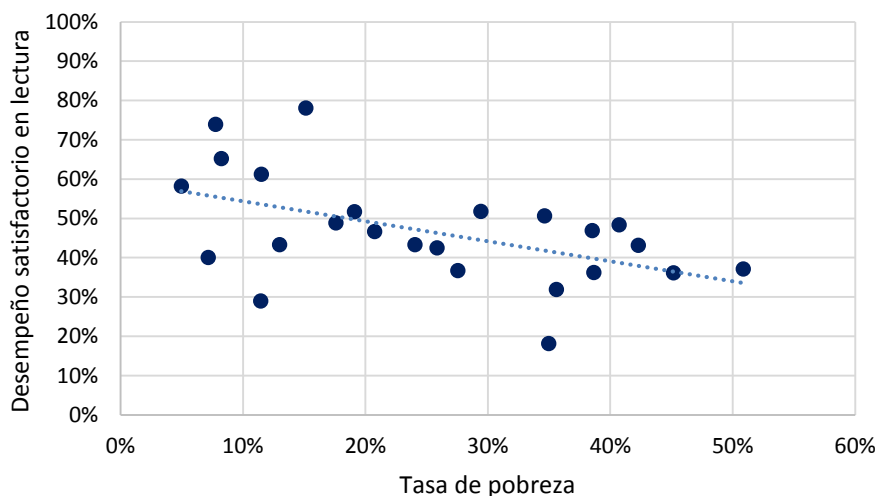


Fuente: ECE y Censo Escolar 2015. Elaboración propia.

Determinantes de demanda: Tasa de pobreza

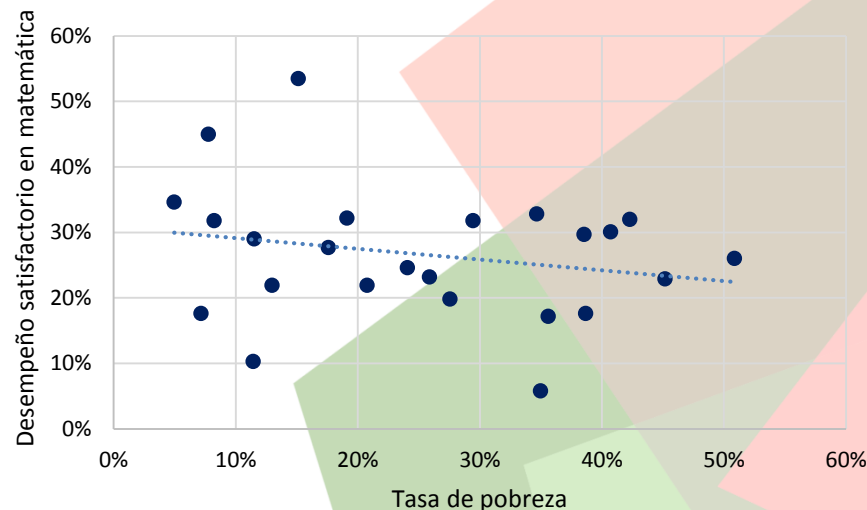
- Parece existir una relación negativa entre el desempeño educativo de los niños y la condición de pobreza:** las regiones que cuentan con menores tasas de pobreza son aquellas donde los alumnos muestran un mejor desempeño en comprensión lectora y matemáticas (2015).

Desempeño satisfactorio en lectura vs. Tasa de pobreza, según regiones 2015



Fuente: ECE y ENAHO 2015. Elaboración propia.

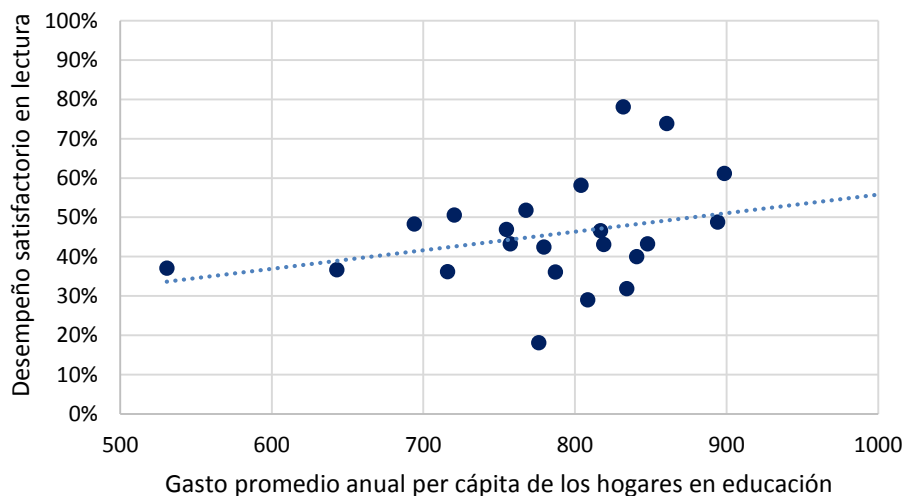
Desempeño satisfactorio en matemática vs. Acceso a servicios básicos, según regiones 2015



Determinantes de demanda: Gasto en educación de los hogares

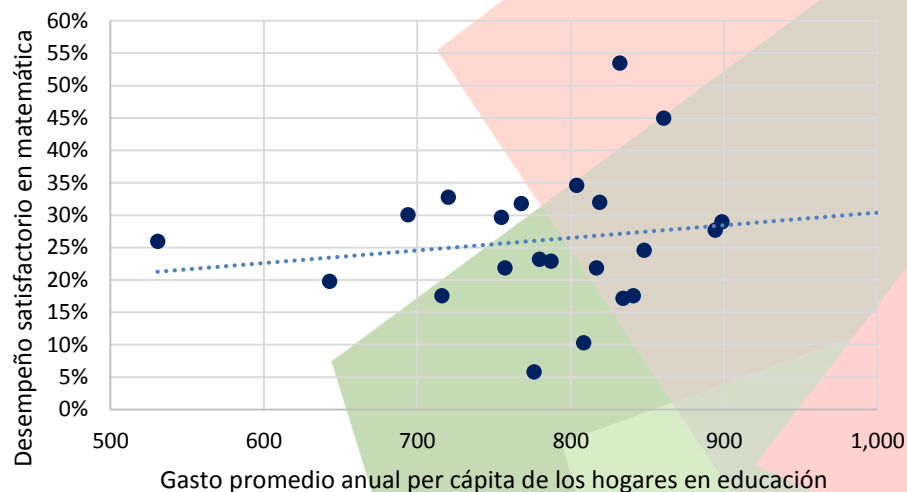
- **Parece existir una relación positiva entre el desempeño educativo de los niños y el gasto del hogar en educación:** las regiones donde los hogares cuyo gasto en educación es mayor son aquellas donde los alumnos muestran un mejor desempeño en comprensión lectora y matemáticas (2015).

**Desempeño satisfactorio en lectura vs.
Gasto del hogar en educación, según regiones 2015**



Fuente: ECE y ENAHO 2015. Elaboración propia.

**Desempeño satisfactorio en matemática vs.
Gasto del hogar en educación, según regiones 2015**



- Para el estudio se construyó una base de datos ad-hoc que permitiera vincular los resultados de los estudiantes de la Evaluación Censal de Estudiantes (ECE) de los años 2013, 2014 y 2015 con los datos del Módulo de Mediciones Antropométricas de la ENAHO (CENAN de los años 2008, 2009, 2010, 2011, 2012). Es decir, se identificó la historia nutricional del niño que rindió la prueba en el año “x” (cuando tenía 7 años), siendo esta historia los indicadores antropométricos registrados cuando el niño tenía 0, 1, 2, 3, 4 o 5 años.
- La ECE es la evaluación que anualmente realiza el Minedu para medir los aprendizajes de todos los estudiantes de segundo grado de primaria del país. Evalúa aprendizajes en matemática y lenguaje y en la muestra los puntajes generados van de 237 a 810 en lenguaje y de 130 a 869 en matemática.
- El CENAN es un módulo antropométrico que excepcionalmente implementó el INEI a las población observada en la muestra de la ENAHO de los años 2008-2012. Reporta indicadores de talla, peso y hemoglobina. De esta base se obtuvo el indicador z-score talla/edad que se utiliza para medir el estado nutricional acumulativo hasta la fecha de toma de la información (también se utiliza para el cálculo de la tasa de desnutrición crónica).
- En acuerdo con el INEI, el Minedu y el CIES, estas bases fueron vinculadas para identificar el estado nutricional en los años 2008-2012 de los niños que rindieron la prueba ECE en los años 2013-2015 generando una base de datos longitudinal única. Los criterios de pegado fueron: (i) nombre similar con una coincidencia superior a 80% en el número de caracteres; (ii) fecha de nacimiento similar con una discrepancia de hasta 1 año; (iii) sexo similar; (iv) igual departamento/provincia y distrito de residencia/asistencia a la escuela.
- El pegado fue realizado por el INEI. El equipo investigador no tuvo acceso a las bases nominadas, sino únicamente a la base resultante innominada. El resultado de la base será puesta a disposición de la comunidad académica.
- En total, la base pegada tuvo alrededor de 4,000 observaciones. Para el estudio, esta base fue depurada (se quitaron observaciones panel de la CENAN, aquellos que no tenían información antropométrica y del rendimiento educativo, y se forzó a tener niños con la misma fecha de nacimiento), por lo cual la base estimable finalmente tuvo alrededor de 2,415 observaciones.

Construyendo conocimiento para mejores políticas

- La estrategia es estimar función de producción empírica de educación a nivel de estudiantes, la cual vincula variables de resultado educativo con un conjunto de insumos (Rivkin y otros, 2005; y Glewwe y Miguel, 2008), algunos *directamente controlables por el hacedor de política* (condiciones de la oferta) y otros *no controlables asociados a las características de las familias y/o las dotaciones iniciales de los alumnos* (como la nutrición). La especificación empírica que particularmente utilizamos es:

$$R_{ijt} = \alpha + HI'_{ijt}\theta_1 + SI'_{ijt}\theta_2 + N'_{ijt-1}\lambda + A_t + \Gamma_t + \Delta_t + \Pi_j + e_{ijt} \quad (1)$$

- donde R_{it} es el resultado en la prueba ECE, HI representa características propias de los hogares o los estudiantes que pueden estar potencialmente correlacionadas con R_{it} , SI son los recursos educativos que el estudiante recibe en el colegio e incluye la infraestructura educativa, los materiales educativos y las características de los docentes; N es el estado nutricional del estudiante en el periodo infantil temprano (se usa el dato más antiguo); A_t son *dummies* que indican el año de nacimiento del niño; Γ_t son *dummies* que indican el año de medición de la CENAN, Δ_t son *dummies* que denotan el año de medición de la ECE y Π_j son efectos fijos a nivel de dominios (costa urbana, costa rural, etc.). Los parámetros $\beta, \theta_1, \theta_2, \lambda$ indican la contribución de cada factor sobre el resultado educativo, siendo este último el de interés en el presente estudio.
- La especificación (1) permite estimar resultados generales. Sin embargo, nos interesan algunos niveles de heterogeneidad, como la influencia diferenciada de la nutrición en espacios urbanos y rurales, entre varones y mujeres, o condición étnica. Por ello, partiendo de la especificación (1), adoptamos la siguiente forma flexible:

$$R_{ijt} = \alpha + HI'_{ijt}\theta_1 + SI'_{ijt}\theta_2 + \sum_{k=1}^K D_k \times N'_{ijt-1}\lambda_k + A_t + \Gamma_t + \Delta_t + \Pi_j + e_{it} \quad (2)$$

- En este caso, D_k son *dummies* que se activan para cada categoría de la dimensión de heterogeneidad a analizar.

- En el estudio, el parámetro de interés es el asociado al estatus nutricional en el período de la infancia temprana, N'_{it-1} , pero es muy probable que esta variable se encuentre correlacionado con variables no observables en e_{it} , conduciendo a que la especificación (1) arroje resultado inconsistentes. Por ello, adicionalmente optamos por implementar una estrategia de variables instrumentales, teniendo como instrumentos los precios de los alimentos en el período anterior a la fecha de medición de la CENAN.
- Para esto, siguiendo a Outes y otros (2011), tomamos como instrumentos a cuatro grupos de alimentos: cereales, carnes, leches y derivados lácteos y tubérculos.
- Operativamente, este método involucró estimar el siguiente modelo de dos etapas: en la primera se modela el estado nutricional de los niños en su infancia temprana como función de los precios (y otros regresores), y en la segunda se modela el rendimiento como función del estado nutricional predicho en la primera etapa.

$$N_{it-1} = \alpha + HI'_{it}\theta_1 + SI'_{it}\theta_2 + \sum_{j=1}^4 P_{jt-1}\sigma_j + A_t + \Gamma_t + \Delta_t + e_{it} \quad (3)$$

$$R_{it} = \alpha + HI'_{it}\theta_1 + SI'_{it}\theta_2 + N'_{it-1}\lambda + A_t + \Gamma_t + \Delta_t + e_{it} \quad (4)$$

Modelo de Mínimos Cuadrados Ordinarios

- Los resultados del modelo de MCO se muestran a continuación. Como se aprecia, la relación entre nutrición (medido por el z-score talla/edad) y el rendimiento académico es directo: bajo la especificación base, un incremento marginal en el z-score talla/edad incrementa los rendimientos en comprensión lectora en 14 puntos y en 18 los de matemática. Tomando los valores promedio muestrales de 571 en comunicación y 562, significa impactos del orden de 2.5% y 3.2%.
- Más aún, los impactos de la nutrición son robustos en nuestras estimaciones a la inclusión de variables potencialmente explicativas de las diferencias observadas: inclusión de controles a nivel de estudiante, de hogar o de la oferta educativa del plantel.

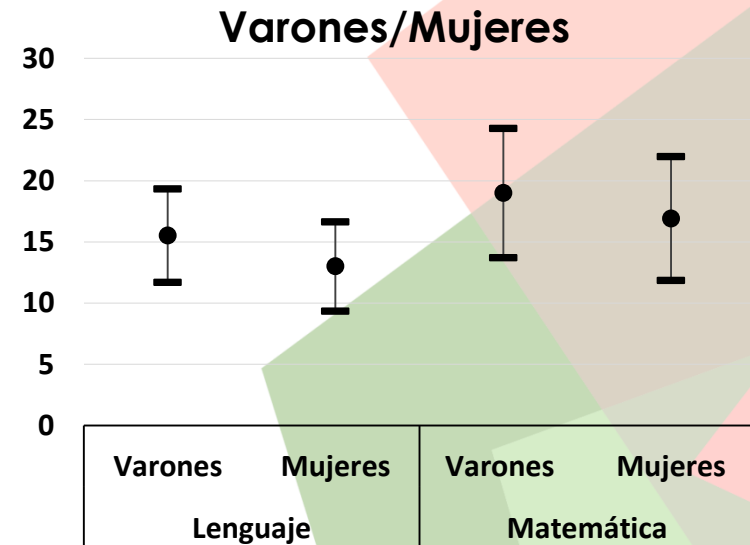
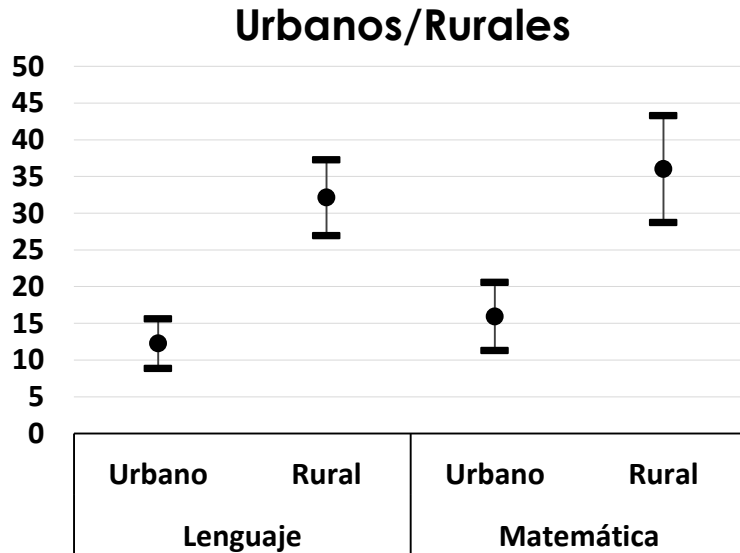
	Comprensión Lectora				Razonamiento Matemático			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Z score - Talla/Edad	14.31*** (1.60) (4.79)	13.53*** (1.59) (4.77)	11.40*** (1.57) (4.88)	10.71*** (1.61) (5.83)	18.00*** (2.21) (6.88)	17.14*** (2.20) (6.85)	14.44*** (2.22) (6.73)	14.02*** (2.28) (7.84)
Constant	483.09*** (25.32)	499.81*** (48.84)	457.29*** (49.09)	431.13*** (49.91)	500.17*** (20.59)	433.11*** (66.02)	394.66*** (65.68)	357.67*** (65.75)
Observations	2,413	2,413	2,413	2,285	2,414	2,414	2,414	2,286
R-squared	0.194	0.212	0.234	0.241	0.120	0.132	0.153	0.162
Controles del estudiante	No	Si	Si	Si	No	Si	Si	Si
Controles del hogar	No	No	Si	Si	No	No	Si	Si
Controles de la oferta	No	No	No	Si	No	No	No	Si

Errores estándar clustereados a nivel de código modular entre paréntesis. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Nota: Todas las regresiones incluyen controles de gestión de la escuela (pública/privada) y ámbito (urbano/rural), además de efectos fijos por año de nacimiento del estudiante, año de medición de la CENAN, año de realización de la ECE y *dummies* por dominio geográfico. Los controles de estudiante son el sexo, la edad en años en momentos de la ECE y si asistió a nivel inicial. Los controles hogar incluyen el nivel educativo del jefe de hogar y su edad. Los controles de oferta incluyen el ratio alumnos/docente, una *dummy* si la escuela tiene simultáneamente agua de red, desagüe de red o letrinas y electricidad; otra *dummy* si tiene biblioteca, y los ratios alumnos/carpetas, alumnos/sillas y alumnos/mesas.

Modelo de Mínimos Cuadrados Ordinarios

- Como ejercicio adicional, se han estimados los parámetros diferenciados para subpoblaciones de interés. Como se aprecia, el impacto es mayor entre estudiantes ubicados en localidades rurales en comparación con aquellos ubicados en localidades urbanas. Sin embargo, no hay diferencias en los impactos a nivel de sexo del estudiante.

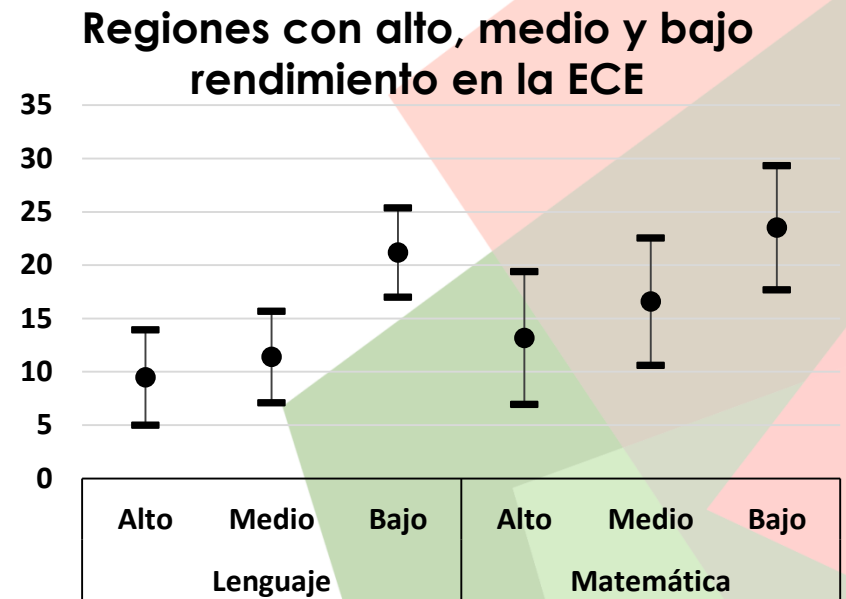
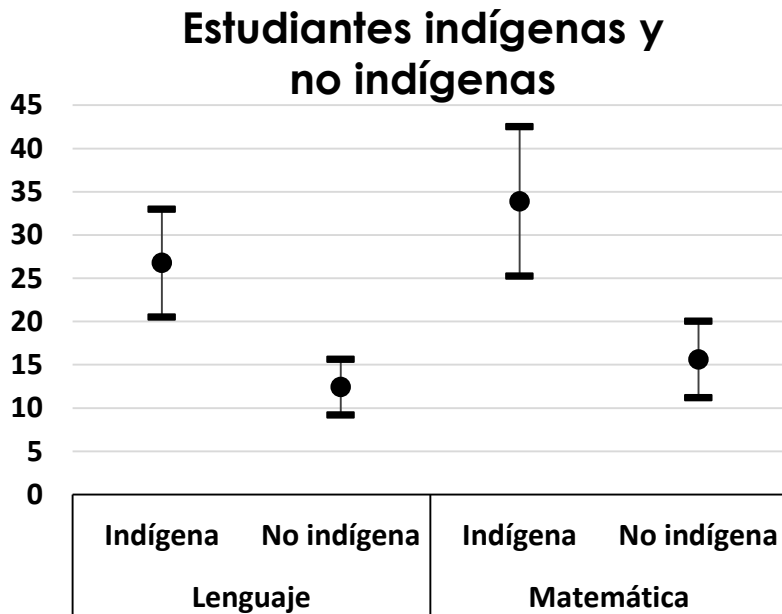


Los gráficos muestran parámetros estimados (puntos negros) e intervalos de confianza al 95%.

Nota: Los modelos se especifican igual que en el caso anterior. Los resultados mostrados corresponden a las versiones (1) y (5)

Modelo de Mínimos Cuadrados Ordinarios

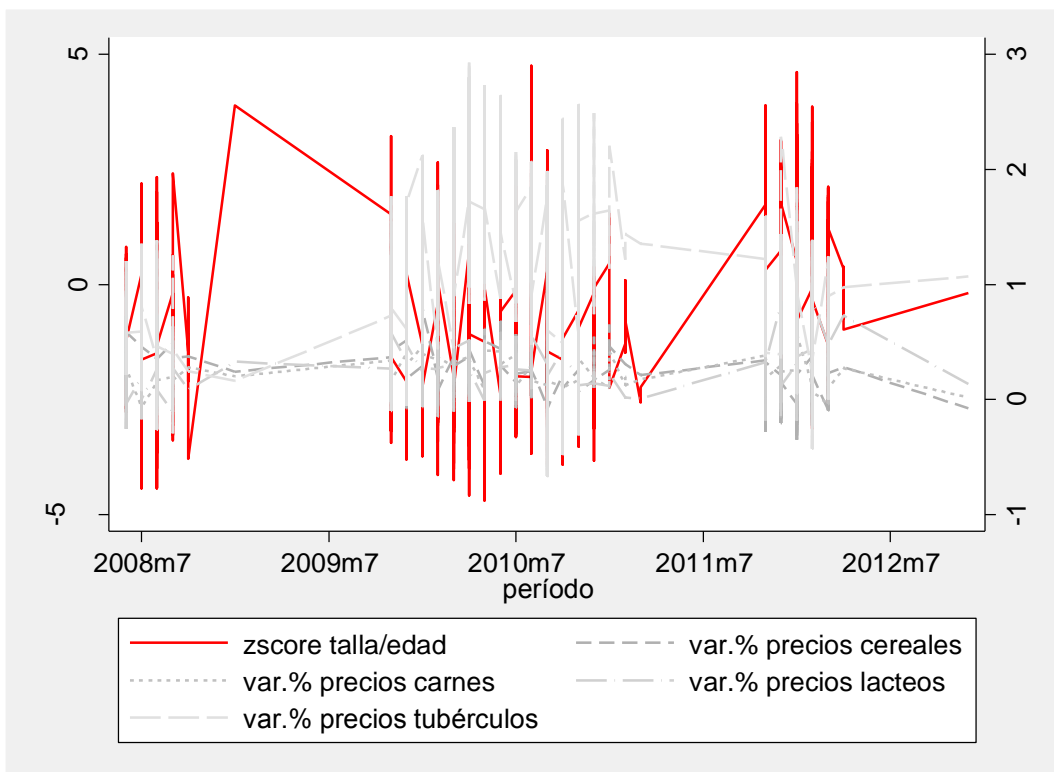
- Adicionalmente, se hicieron estimaciones para estudiantes indígenas (con lengua materna quechua, aymara o alguna lengua amazónica) y no indígenas (con lengua materna castellano). Aquí, también, los resultados son más ventajosos para estudiantes indígenas, en particular en el matemática.
- Finalmente, se han estimados impactos heterogéneos entre estudiantes que residen en regiones con alto, medio y bajo rendimiento en la ECE. En este punto se observan diferencias significativas entre los grupos, pues parece que la nutrición incrementa más los rendimientos en las regiones con peores resultados agregados.



Los gráficos muestran parámetros estimados (puntos negros) e intervalos de confianza al 95%.

Nota: Los modelos se especifican igual que en el caso anterior. Los resultados mostrados corresponden a las versiones (1) y (5)

Modelo de Variables Instrumentales



- En rojo se encuentra los estimados del z-score talla/edad para los estudiantes de la muestra. En plomo las inflaciones promedio en tres años previos a la fecha de medición del indicador. Al igual que Outes y otros (2011) encontramos una correlación negativa entre ambas variables.
- Es decir, la inflación de alimentos cuando el niño está en la etapa crítica de su desarrollo infantil temprano afecta negativamente su status nutricional debido a las familias tendrán mayores dificultades para alimentarlos adecuadamente
- Asumimos que la inflación de alimentos no se correlaciona con el rendimiento académico de forma directa por lo que se puede considerar exógeno

Modelo de Variables Instrumentales

- Resultados de primera etapa indican presencia de instrumentos débiles.

	(1)	(2)	(3)		(1)	(2)	(3)
VARIABLES	haz06	haz06	haz06	VARIABLES	haz06	haz06	haz06
Grupo 1 x Edad=1	0.93*** (0.31)	0.95*** (0.31)	0.93*** (0.31)	Grupo 3 x Edad=1	-0.82* (0.46)	-0.89** (0.46)	-0.74 (0.48)
Grupo 1 x Edad=2	0.04 (0.28)	0.03 (0.27)	0.07 (0.28)	Grupo 3 x Edad=2	-0.39 (0.33)	-0.36 (0.32)	-0.35 (0.33)
Grupo 1 x Edad=3	-0.16 (0.25)	-0.17 (0.25)	-0.24 (0.24)	Grupo 3 x Edad=3	0.69*** (0.23)	0.57** (0.23)	0.55** (0.24)
Grupo 1 x Edad=4	0.06 (0.24)	0.07 (0.24)	-0.02 (0.25)	Grupo 3 x Edad=4	-0.40* (0.23)	-0.34 (0.23)	-0.36 (0.23)
Grupo 1 x Edad=5	-1.38*** (0.46)	-1.22*** (0.47)	-1.38*** (0.46)	Grupo 3 x Edad=5	-0.73*** (0.25)	-0.65** (0.26)	-0.66** (0.27)
Grupo 2 x Edad=1	0.96** (0.46)	1.04** (0.45)	0.94** (0.47)	Grupo 4 x Edad=1	-0.07 (0.15)	-0.04 (0.15)	-0.03 (0.17)
Grupo 2 x Edad=2	0.79** (0.34)	0.72** (0.33)	0.68* (0.35)	Grupo 4 x Edad=2	-0.14 (0.10)	-0.10 (0.10)	-0.11 (0.10)
Grupo 2 x Edad=3	-0.20 (0.31)	-0.15 (0.31)	0.06 (0.31)	Grupo 4 x Edad=3	-0.04 (0.07)	-0.01 (0.07)	-0.03 (0.07)
Grupo 2 x Edad=4	-0.03 (0.35)	-0.16 (0.35)	-0.01 (0.35)	Grupo 4 x Edad=4	0.01 (0.08)	0.03 (0.08)	0.03 (0.08)
Grupo 2 x Edad=5	1.17** (0.55)	1.03* (0.56)	1.19** (0.57)	Grupo 4 x Edad=5	-0.15 (0.11)	-0.16 (0.11)	-0.16 (0.11)
				Observations	2,414	2,414	2,286
				R-squared	0.437	0.451	0.454
				F-stat	42.1	43.0	36.9
				Cragg y Donald	3.9	3.9	3.3

Errores estándar clustereados a nivel de código modular entre paréntesis. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Todas las regresiones incluyen controles de gestión de la escuela (pública/privada) y ámbito (urbano/rural), además de efectos fijos por año de nacimiento del estudiante, año de medición de la CENAN, año de realización de la ECE y *dummies* por dominio geográfico. (2) incluye controles de estudiante como el sexo, la edad en años en momentos de la ECE y si asistió a nivel inicial y los controles hogar como el nivel educativo del jefe de hogar y su edad. (3) Incluye además los controles de oferta como el ratio alumnos/docente, una *dummy* si la escuela tiene simultáneamente agua de red, desagüe de red o letrinas y electricidad, si tiene biblioteca, y los ratios alumnos/carpetas, alumnos/sillas y alumnos/mesas.

Los valores críticos de Stock y Yogo (2005) son 21.38 para un sesgo relativo máximo de VI de 5%; 11.45 para un sesgo relativo máximo de 10%; 6.28 para un sesgo relativo máximo de 20%, y 4.48 para un sesgo relativo máximo de 30%

Modelo de Variables Instrumentales

- Implementamos la segunda etapa del estimador de variables instrumentales usando el método 2SLS. Además, se realizaron test de exogeneidad de Sargant y test de Anderson-Rubin para inferencia en presencia de instrumentos débiles. Un incremento marginal en el z-score talla/edad incrementa los rendimientos en comprensión lectora en [20 26] puntos y en [20 27] los de matemática. Esto significa impactos del orden de 6% y 4%, respectivamente.

	Comunicación			Matemática		
	(1)	(2)	(3)	(1)	(2)	(3)
Z score - Talla/Edad	26.35*** (8.86)	19.50** (9.44)	21.49** (9.42)	26.94** (12.77)	20.35 (13.93)	21.96 (13.63)
Constant	506.88*** (28.21)	470.79*** (48.74)	445.78*** (49.86)	516.80*** (27.15)	411.07*** (64.47)	375.16*** (64.65)
Observations	2,413	2,413	2,285	2,414	2,414	2,286
R-squared	0.174	0.225	0.226	0.113	0.149	0.157
Test de Sargan (χ^2)	25.39	25.40	22.34	26.39	26.61	28.03
Test de Sargan [$P > \chi^2$]	0.15	0.15	0.27	0.12	0.11	0.08
Test Anderson-Rubin (F)	1.657	1.444	1.373	1.544	1.450	1.551
Test Anderson-Rubin P[F = 0]	0.03	0.09	0.12	0.06	0.09	0.06
Controles del estudiante	No	Si	Si	No	Si	Si
Controles del hogar	No	Si	Si	No	Si	Si
Controles de la oferta	No	No	Si	No	No	Si

Errores estándar clustereados a nivel de código modular entre paréntesis. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Nota: Los regresores incluidos en cada caso son similares a los comentados en el cuadro de la lámina anterior.

Modelo de Variables Instrumentales

- Como segundo ejercicio de robustez, se corrieron las regresiones para diferentes métodos de estimación a los de la tabla anterior. En particular, se implementó la estimación por el método de máxima verosimilitud con información limitada (LIML) y método generalizado de momentos (GMM). Estos métodos son robustos a la presencia de instrumentos débiles, pero a costa de menor eficiencia. Encontramos que los efectos (LIML) se ubican en el orden de 34 (comunicación) y 32 (matemática) lo que significa efectos del orden de alrededor del 6% en ambos casos.

	Comunicación			Matemática		
	GMM	2SLS	LIML	GMM	2SLS	LIML
Z score - Talla/Edad	26.35*** (8.86)	27.37*** (9.17)	34.23** (14.05)	26.94** (12.77)	27.46** (13.06)	32.38 (19.76)
Constante	506.88*** (28.21)	500.28*** (29.16)	509.30*** (33.06)	516.80*** (27.15)	512.62*** (27.62)	519.10*** (34.27)
Observations	2,413	2,413	2,413	2,414	2,414	2,414
R-squared	0.174	0.171	0.140	0.113	0.113	0.104
Test de Sargan (χ^2)	25.39	25.39	24.80	26.39	26.39	26.26
Test de Sargan [$P > \chi^2$]	0.15	0.15	0.17	0.12	0.12	0.12
Test Anderson-Rubin (F)	1.657	1.657	1.657	1.544	1.544	1.544
Test Anderson-Rubin $P[F=0]$	0.03	0.03	0.03	0.06	0.06	0.06

Errores estándar clustereados a nivel de código modular entre paréntesis. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Nota: Las regresiones están especificadas como (1) en la lámina anterior.

Hallazgos preliminares

- Los niveles de nutrición en el periodo infantil temprano afectan positiva y significativamente los resultados académicos logrados en la niñez. Esta hipótesis se ha verificado usando datos niños peruanos a partir de los indicadores z-score (talla / edad) en el periodo 0 – 5 años y los resultados de una prueba estandarizada tomada en segundo de primaria (7 años). Los resultados son robustos a la inclusión de variables de oferta y demanda estándares en la literatura y se mantienen bajo supuestos de endogeneidad de la variable de interés.
- En particular se encuentra que una innovación de una 1 desviación estándar en el indicador nutricional genera un efecto positivo de alrededor de 30 puntos en los resultados de las pruebas tanto de razonamiento verbal como matemático. Esto significa que el cambio marginal experimentado en la nutrición mejora en 6% los resultados académicos. Si utilizamos estos resultados para simular linealmente lo ocurrido entre 2013 y 2015 (usando datos muestrales) encontramos que el factor nutricional sería responsable de más del 50% del cambio experimentado en la mejora educativa reciente.
- Los ejercicios de correlación adicionales presentados en el documento sugieren además que los resultados tienden a ser mayores entre los estudiantes ubicados en zonas rurales, que pertenecen a poblaciones indígenas y aquellos con un nivel de rendimiento inicial menor. Es decir, la nutrición tiende a ser una condición más crítica en el rendimiento académico para aquellos con características socioeconómicas menos favorables o que presenten mayores rezagos en su desarrollo educativo. De este modo, identificamos espacios de articulación entre políticas de salud en el periodo infantil temprano y la etapa educativa en la niñez que permitan sostener las mejoras recientes en este último indicador y que potencialmente permitan acelerar el desarrollo de los grupos menos favorecidos.
- Trabajo futuro: efectos fijos a nivel de familia, incluir variables de demanda adicionales, estimaciones para las variables de oferta, explorar instrumentos adicionales (clima) y desarrollar las recomendaciones de política