

¿VALE LA PENA ESTUDIAR? RETORNOS DE LA EDUCACIÓN EN EL PERÚ CON EVIDENCIA DE DATOS DE PANEL

Juliana Mery Bautista Lopez. jbautistal@unsa.edu.pe. Universidad Nacional de San Agustín de Arequipa. <https://orcid.org/0000-0003-3359-3618>

Hebert Suarez Cahuana. hsuarezc@unsa.edu.pe. Universidad Nacional de San Agustín de Arequipa. <https://orcid.org/0000-0003-2433-6804>

DOI: <https://doi.org/10.71727/rae.v4i2.350>

RESUMEN

Objetivo: Este estudio busca estimar los retornos de la educación sobre el ingreso por hora en el Perú (2018-2022), controlando por heterogeneidad no observable y posibles sesgos por variable omitida. **Metodología:** Se aplica un modelo de datos de panel (ENAHO) con especificación Mincer, comparando efectos fijos y aleatorios. Se valida la consistencia del modelo mediante la prueba de Hausman. **Resultados:** El modelo de efectos fijos es más consistente, mostrando que los retornos educativos son menores que en estimaciones MCO tradicionales, lo que sugiere un sesgo por variables omitidas en enfoques no panel. **Conclusión:** Los hallazgos confirman la importancia de usar datos de panel para estimaciones precisas, revelando que estudios previos podrían haber sobrestimado el impacto de la educación en los ingresos.

Palabras clave: Retornos educativos, efectos fijos, sesgo de estimación.

ABSTRACT

Objective: This study aims to estimate the returns to education on hourly income in Peru (2018–2022), controlling for unobserved heterogeneity and potential omitted variable bias. **Methodology:** A panel data model (ENAHO) with a Mincer specification is applied, comparing fixed and random effects. Model consistency is validated using the Hausman test. **Results:** The fixed-effects model proves more consistent, showing that educational returns are lower than in traditional OLS estimates, suggesting omitted variable bias in non-panel approaches. **Conclusion:** The findings confirm the importance of using panel data for accurate estimations, revealing that previous studies may have overestimated the impact of education on income.

Keywords: Returns to education, fixed effects, estimation bias.

INTRODUCCIÓN

La relación entre educación e ingresos en el Perú sigue siendo un tema de debate en la economía laboral. Como señalan Samuelson y Nordhaus (2010), una fuerza laboral más calificada impulsa la productividad nacional, mientras que a nivel individual, mayores niveles educativos suelen traducirse en mejores ingresos (Becker, 1962). Sin em-

bargo, persiste un debate crucial: ¿Las estimaciones convencionales de estos retornos están distorsionadas por no considerar factores no observables como la habilidad innata? Esta investigación aborda este problema. En particular, una pregunta clave es si los métodos tradicionales para estimar los retornos de la educación —basados en datos de corte transversal— están sobreestimando el verdadero efecto debido a sesgos por variables omitidas, como la habilidad innata o la motivación personal (Griliches, 1975). Esta investigación busca aportar evidencia rigurosa al estimar el impacto de los años de escolaridad sobre el ingreso por hora, utilizando los datos de panel de la ENAHO (2018-2022) y comparando los modelos de efectos fijos y aleatorios.

La respuesta a esta pregunta tiene implicancias directas en políticas públicas. Si los retornos educativos han sido históricamente sobreestimados, las decisiones de inversión en educación —tanto de individuos como del Estado— podrían estar basadas en expectativas poco realistas. Por ejemplo, el cierre de universidades de baja calidad (Ley 30220) y las mediciones de eficiencia del gasto universitario (MEF, 2019) requieren de estimaciones precisas sobre cuánto realmente aporta la educación a los ingresos.

El concepto de capital humano (Becker, 1962) sostiene que la educación incrementa la productividad laboral, pero su medición enfrenta un desafío metodológico: ¿cómo aislar el efecto puro de los años de estudio, controlando por características no observables como la ambición o la disciplina? Estudios peruanos previos (Yamada, 2006; Barboza, 2019) han usado mínimos cuadrados ordinarios (MCO) o variables instrumentales, pero ninguno ha aplicado sistemáticamente datos de panel, que permiten controlar la heterogeneidad individual no observable (Wooldridge, 2010).

La literatura internacional ha demostrado que ignorar estos factores lleva a sesgos. Por ejemplo, Ashenfelter y Krueger (1994), al estudiar gemelos, encontraron que las estimaciones tradicionales exageraban los retornos educativos en hasta un 30%.

En el Perú, los estudios relacionados revelan una evolución interesante. Mientras Cangalaya (1973) encontró retornos del 10.5% para educación superior, investigaciones posteriores como las de (Rodríguez, 1993; Yamada, 2006; Yamada, 2016) mostraron variaciones significativas según nivel educativo y tipo de institución. Particularmente revelador es el hallazgo de Khandker (1990) sobre mayores retornos en universidades privadas, que plantea interrogantes sobre calidad educativa y mecanismos de selección. Estos estudios, sin embargo, comparten la limitación de usar datos trans-

versales que, como demostró Heckman et al. (2008), pueden producir estimadores inconsistentes.

En este contexto, esta investigación plantea tres contribuciones clave: Metodológica: Al aplicar modelos de panel, se mitiga el sesgo por variables omitidas, ofreciendo estimaciones más confiables. Empírica: Se actualiza la evidencia para el periodo 2018-2022, crucial por los cambios en el mercado laboral postpandemia. De política pública: Los resultados pueden contribuir con la implementación de reformas educativas y plataformas como “Ponte en Carrera” (MINEDU & MTPE, 2023), que guían a los estudiantes en sus decisiones sobre qué carrera elegir.

Como señalan Musgrave y Musgrave (1992), los beneficios sociales de la educación van más allá de los ingresos individuales, incluyendo externalidades positivas para la democracia y la ciudadanía y una vida más satisfactoria derivada de una mejor educación. Sin embargo, medir adecuadamente el componente económico sigue siendo esencial para la toma de decisiones, especialmente en un contexto de recursos limitados.

La discusión teórica no está exenta de controversias. Mientras algunos sostienen que la educación actúa principalmente como señal de productividad (Stiglitz, 1975; Mankiw, 2012)), otros enfatizan su rol en la formación de habilidades (Mincer, 1974). Este estudio no resuelve el debate, pero aporta evidencia sólida para situarlo en la realidad peruana.

Con este fin, el artículo se organiza de la siguiente manera: Se detalla la metodología, luego se presentan los resultados de los retornos estimados y se contrastan con los obtenidos por MCO, validando los hallazgos con la prueba de Hausman. Luego la discusión y conclusiones: Se analizan las implicancias para políticas públicas y se finalmente se proponen líneas futuras de investigación.

METODOLOGÍA

Este estudio emplea un enfoque cuantitativo con diseño no experimental longitudinal, de nivel explicativo (Hernandez-Sampieri y Mendoza, 2018), utilizando datos de panel de la ENAHO 2018-2022 (INEI, 2023). El modelo base sigue la ecuación de Mincer (1974), donde el logaritmo del ingreso por hora (variable dependiente) se relaciona con los años de educación (variable independiente) y variables de control como edad, sexo, estado civil y características laborales.

Tipo de investigación: Explicativa, buscando establecer relaciones causales entre educación e ingresos

Fuente de datos: Encuesta Nacional de Hogares (ENAH) con diseño muestral panel

Muestra: 3,916 individuos de la PEA ocupada, observados durante 5 años (n=19,580 observaciones)

VARIABLES CLAVE: Dependiente: Ingreso por hora (logaritmo)

Independiente: Años de educación completados

Control: Edad, sexo, experiencia, sector económico, etc.

Técnicas analíticas: Modelos de efectos fijos para controlar heterogeneidad no observable
Pruebas de robustez y especificación (Hausman)
Errores estándar agrupados (clúster) para controlar la correlación serial y heteroscedasticidad.

datos a formato long y aplicando el factor de expansión muestral para garantizar representatividad.

Resultados

Características de la Muestra

El estudio utilizó datos de la Encuesta Nacional de Hogares (ENAH) en formato panel para los años 2018-2022, con un diseño muestral que permite seguir a los mismos individuos a lo largo del tiempo. La muestra final incluyó 3,476 individuos con un promedio de 3.68 observaciones por persona, totalizando 12,781 observaciones.

Estimación de los Retornos Educativos

Modelo de Efectos Fijos: El modelo controla por heterogeneidad no observable (ej. habilidad innata) y arroja estimaciones más precisas que los métodos tradicionales. El modelo a estimar es:

$$\text{Log(Ingreso hora)} = \beta_1 + \beta_{educ}educ + \dots + \nu_i + \epsilon_{it}$$

Donde:

- β_{educ} : es el retorno de un año adicional de educación.
- ν_i : es el efecto fijo individual no observable invariante en el tiempo (habilidad innata)
- ϵ_{it} : Es el término del error variante en el tiempo e individuo.
- $educ$: Años de educación del individuo.
- $\mu_i = \nu_i + \epsilon_{it}$ es el error compuesto a nivel de individuo.

La metodología cumple con los criterios de investigación económica al:

- Plantear relaciones causales teóricamente fundamentadas
- Utilizar técnicas econométricas adecuadas para datos de panel
- Controlar por variables omitidas mediante efectos fijos
- Permitir inferencias válidas a nivel nacional

El análisis se realizó con Stata 18, transformando los

El análisis con efectos fijos confirma que cada año adicional de educación incrementa los ingresos en

Tabla 1

Análisis de regresión de los efectos fijos de los años de educación sobre el ingreso por hora 2018-2022

Variable dependiente: Logaritmo del ingreso por hora								
Regresor	(1)	(2)	(3)	(4)	(5)	(6)	(7)	
Educación (años)	0.087 (0.0022) [0.082, 0.091] 0.000	0.028 (0.0081) [0.012, 0.044] 0.001	0.027 (0.0082) [0.011, 0.043] 0.001	0.028 (0.0083) [0.012, 0.044] 0.001	0.025 (0.0082) [0.009, 0.041] 0.002	0.022 (0.0081) [0.007, 0.038] 0.006	0.02 (0.0138) [-0.007, 0.047] 0.14	
Hombre=1				-0.021 (0.058) [0.135, 0.093] 0.714	0.262 (0.164) [0.060, 0.584] 0.111	0.286 (0.164) [0.037, 0.608] 0.082		
Edad (años)				0.028 (0.017) [0.006, 0.062] 0.101	0.026 (0.017) [0.007, 0.060] 0.121	0.024 (0.017) [0.010, 0.058] 0.165	0.052 (0.037) [-0.021, 0.125] 0.163	

Casado=1				0.042 (0.07) [0.096, 0.179] 0.553	0.053 (0.071) [0.087, 0.193] 0.458	0.059 (0.072) [0.082, 0.199] 0.411	0.131 (0.097) [-0.059, 0.320] 0.176
Antigüedad				0.007 (0.002) [0.002, 0.011] 0.006	0.006 (0.002) [0.002, 0.011] 0.006	0.006 (0.002) [0.002, 0.011] 0.005	0.005 (0.004) [-0.003, 0.014] 0.218
Covid				-0.059 (0.061) [0.178, 0.06] 0.335	.	.	.
Años	2018-2022	2018-2022	2018-2022	2018-2022	2018-2022	2018-2022	2018 y 2022
¿Efectos fijos?	No	Si	Si	Si	Si	Si	Si
¿Efectos temporales?	No	No	Si	Si	Si	Si	Si
¿Tipo de actividad?	No	No	No	No	Si	Si	Si
¿Tamaño de empresa?	No	No	No	No	No	Si	Si
Estadísticos F y sus valores p para las pruebas de exclusión de grupos de variables							
Efectos temporales=0		18.65 (<0.001)	0.084 (0.9684)	3.99 (0.0031)	3.85 (0.0040)	0.20 (0.6544)	
Actividades=0				5.23 (<0.001)	3.10 (<0.001)	3.83 (<0.001)	
Tamaños=0					8.61 (<0.001)	3.54 (<0.0034)	
$H_0 : u_i = 0$		4.8425 (<0.001)	4.9803 (<0.001)	4.5702 (<0.001)	4.1860 (<0.001)	3.9167 (<0.001)	2.3889 (<0.001)
$N \times T$	12 474	12 474	12 474	11 477	11 477	11 477	5 148
\bar{R}^2	0.193	0.002	0.017	0.016	0.035	0.042	0.062

Nota: Estas regresiones fueron estimadas con la ENAHO panel 2018-2022. Las regresiones (1) a (6) utilizan todos los años, (7) utiliza sólo 2018 y 2022. Los errores estándar están entre paréntesis, los intervalos de confianza al 95% y el valor p están bajo los coeficientes y en los casos de las pruebas conjunta el valor p se presenta debajo del valor F. Se aplicó el factor de expansión.

aproximadamente 2.2%, mostrando una relación positiva y significativa. Se detectan efectos temporales relevantes en todas las especificaciones, excepto en el modelo (4). Las pruebas estadísticas rechazan consistentemente la hipótesis nula $H_0 : u_i = 0$, validando la existencia de efectos individuales no observables a nivel de individuo $v_i \neq 0$ y justificando el uso de modelos de panel frente a MCO. La validez del modelo de efectos fijos depende crucialmente de la correlación entre habilidad innata y educación, supuesto plausible dado que personas más hábiles tienden a obtener mayor escolaridad. Solo si esta correlación no existiera - caso no respaldado por los datos - el modelo de efectos aleatorios sería preferible por su mayor eficiencia estadística. Estos resultados destacan la importancia de controlar por heterogeneidad no observable en estudios de retornos educativos.

PRUEBA DE HAUSMAN

Los resultados de la prueba de Hausman (χ^2), donde en todos los modelos se rechaza la hipótesis nula ($p < 0.05$), confirmando que el estimador de efectos fijos es consistente y preferible sobre efectos aleatorios

para este análisis. Esto valida las estimaciones reportadas.

Tabla 2

Estadístico χ^2 para la prueba de Hausman para efectos fijos y aleatorios.

Estadístico	Modelo					
	(2)	(3)	(4)	(5)	(6)	(7)
χ^2	48.12	81.72	73.85	116.82	114.17	44.47
Valor p	(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)

Conclusiones

Este estudio buscó resolver una pregunta central en economía laboral: ¿Cuál es el verdadero impacto de un año adicional de educación sobre los ingresos en el Perú, una vez controlado por factores no observables como la habilidad innata? A través de un análisis pionero con datos de panel de la ENAHO (2018-2022), se superaron las limitaciones de estudios previos que usaban cortes transversales y enfrentaban sesgos por variables omitidas.

Los resultados revelaron: Un retorno educativo real del ~2% por año adicional de escolaridad (modelo de efectos fijos), significativamente menor que las estimaciones tradicionales (8.7% con MCO). Esta diferencia confirma que el 75% del retorno aparente en estudios previos se explica por características individuales no observables (ej. motivación, disciplina).

Los efectos temporales son significativos, lo que sugiere que factores macroeconómicos (ej. crisis por COVID-19) influyen en la relación educación-ingresos. El modelo de efectos fijos es superior al de efectos aleatorios, según la prueba de Hausman ($p < 0.001$ en todas las especificaciones), validando su uso para estimaciones causales.

Los retornos más bajos cuestionan la eficiencia marginal del gasto en educación superior y resaltan la necesidad de priorizar calidad educativa sobre cobertura. Herramientas como "Ponte en Carrera" (MINEDU) deberían ajustar sus expectativas salariales para reflejar retornos reales, no sesgados. Los hallazgos respaldan la teoría del capital humano (Becker, 1962), pero con retornos menores a los esperados, lo que podría indicar problemas de calidad educativa o sobreoferta de profesionales.

Algunas limitaciones: Variabilidad temporal limitada: Solo 5 años de panel impiden analizar efectos a largo plazo. No se captura el "efecto diploma": La medición lineal de años de educación no distingue entre completar ciclos educativos (ej. terminar la universidad vs. dejar estudios). Muestra restringida: Exclusivamente población ocupada, lo que podría excluir dinámicas del desempleo.

Para futuras líneas de investigación se sugiere extender el panel a más años para analizar dinámicas intertemporales. Incorporar variables de calidad educativa: Ranking de universidades, resultados de licenciamiento (SUNEDU). Explorar heterogeneidades como: Diferencias por género, región o tipo de institución (pública vs. privada). Incluir mecanismos no monetarios (Satisfacción laboral o movilidad social como retornos de la educación).

Este estudio demuestra que, en el Perú, la educación sigue siendo un motor de movilidad económica, pero con retornos más modestos de lo que se creía. Los resultados exigen un cambio de paradigma: más allá de acumular años de escolaridad, el país debe garantizar que cada año educativo agregue habilidades relevantes para el mercado laboral. Como señaló Mincer (1974), "la educación es inversión, no gasto", pero su valor depende críticamente de qué y cómo se aprende.

Urge integrar estos hallazgos en el diseño de políticas educativas y laborales, combinando expansión de cobertura con mecanismos de aseguramiento de calidad. Solo así se maximizará el impacto del capital humano en el desarrollo peruano.

REFERENCIAS

- Ashenfelter, O., & Krueger, A. (1994). Estimates of the Economic Return to Schooling from a New Sample of Twins. *The American Economic Review*, 84(5), 1157–1173.
- Barboza, G. (2019). Retornos de la educación superior universitaria en el mercado laboral de las regiones de Arequipa, Moquegua, Puno y Tacna: 2007-2017. Universidad Nacional del Altiplano.
- Becker, G. (1962). Investment in Human Capital: A Theoretical Analysis. *Journal of Political Economy*, 70(5), 9–49.
- Becker, G. (2008). Human Capital. In *The Concise Encyclopedia* (pp. 248–256).
- Cangalaya, L. (1973). Rentabilidad de la inversión en el Sistema Educativo Nacional. Escuela Superior de Administración Pública.
- Griliches, Z. (1975). Estimating the Returns to Schooling: Some Econometrics Problems. *Econometrica*, 45(1), 1–22.
- Heckman, J. J., Lochner, L. J., & Todd, P. E. (2008). Earnings Functions and Rates of Return. *Journal of Human Capital*, 2(1), 1–31. <https://doi.org/10.1086/587037>
- Hernandez-Sampieri, R., & Mendoza, C. (2018). Metodología de la investigación. Las rutas cuantitativa, cualitativa y mixta. (1st ed.). McGraw-Hill.
- INEI. (2023). Ficha técnica de la base de datos panel 2018-2022.
- Khandker, S. R. (1990). Labor market participation, returns to education, and male-female wage differences in Peru (Vol. 461). World Bank Publications.
- Mankiw, G. (2012). Principios de Economía (6th ed.). Cengage Learning.
- MEF. (2019). Calendario de evaluaciones independientes para el año fiscal 2019.
- Mincer, J. (1974). Schooling, experience and earnings. National Bureau of Economic Research.
- MINEDU, & MTPE. (2023). Ponte en carrera.
- Musgrave, R., & Musgrave, P. (1992). Hacienda pública. Teórica y aplicada (5th ed.). McGraw-Hill/Interamericana de España, S. A. U.
- Rodríguez, J. (1993). Retornos económicos de la educación en el Perú. 112.
- Samuelson, P., & Nordhaus, W. (2010). Economía con aplicaciones a Latinoamérica (19th ed.).
- Stiglitz, J. E. (1975). The Theory of "Screening," Education, and the Distribution of Income. *The American*

- can *Economic Review*, 65(3), 283–300.
- Suarez Cahuana, H. (2024). Estimación de los retornos de la educación sobre el ingreso: una aplicación de datos de panel en el Perú, 2018 – 2022. [Tesis de Pregrado, Universidad Nacional de San Agustín]. Repositorio Institucional de la UNSA. <https://hdl.handle.net/20.500.12773/17848>
- Wooldridge, J. (2010). *Introducción a la econometría* (4th ed.). Cengage Learning.
- Yamada, G. (2006). *Retornos a la educación superior en el mercado laboral*. CIUP Universidad del Pacífico.
- Yamada, G., Lavado, P., & Oviedo, N. (2016). La evidencia de rendimientos de la educación superior a partir de “Ponte en Carrera.” CIUP Universidad del Pacífico.