

Evaluación del Impacto del Programa de Transferencias Condicionadas Juntos mediante Inferencia Causal: Un Análisis con Datos ENAHO 2022

[1]
[1]

¹Universidad Nacional del Altiplano Puno, Estadística e Informática
{

12 de agosto de 2025

Resumen

Este estudio implementa un marco avanzado de inferencia causal que combina Propensity Score Matching (PSM) con Diferencias en Diferencias (DiD) para evaluar el impacto del programa de transferencias monetarias condicionadas “Juntos” del Perú en indicadores multidimensionales de bienestar infantil. Utilizando datos de la ENAHO 2022 que comprenden 5,000 observaciones de hogares, desarrollamos un diseño cuasi-experimental que aborda los desafíos de sesgo de selección y heterogeneidad no observada inherentes a las intervenciones sociales no aleatorizadas. La metodología incorporó regresión logística para la estimación del propensity score (pseudo- $R^2=0.342$, AUC=0.821), matching del vecino más cercano con caliper de 0.05 logrando una tasa de emparejamiento del 82.5 %, y estimación DiD basada en bootstrap con 500 replicaciones para errores estándar robustos. Los resultados demuestran impactos del programa estadísticamente significativos: incremento del 15.2 % en asistencia escolar (IC 95 %: [10.5 %, 19.9 %], $p<0.001$), mejora del 20.3 % en utilización de controles de salud (IC 95 %: [14.2 %, 26.4 %], $p<0.001$), aumento de 35.7 soles mensuales en gasto educativo (IC 95 %: [7.87, 63.53], $p=0.012$), y reducción del 9.8 % en prevalencia de desnutrición crónica (IC 95 %: [-13.3 %, -6.3 %], $p<0.001$). Todas las covariables lograron balance satisfactorio post-matching ($|SMD|$

$i(0,1)$,

validando la estrategia de identificación. Los hallazgos proporcionan evidencia empírica robusta que respalda la efectividad de las transferencias condicionadas para romper los ciclos intergeneracionales de pobreza, con implicaciones para el escalamiento y optimización de programas de protección social en países en desarrollo. Esta investigación contribuye a la literatura creciente sobre metodologías de evaluación de impacto al demostrar la aplicación exitosa de técnicas cuasi-experimentales combinadas en contextos socioeconómicos complejos.

Palabras clave: Inferencia causal, Propensity Score Matching, Diferencias en Diferencias, Transferencias condicionadas, Programa Juntos, ENAHO, Evaluación de impacto, Diseño cuasi-experimental, Inferencia bootstrap, Protección social

1. Introducción

El crecimiento exponencial de los programas de transferencias monetarias condicionadas (TMC) ha transformado fundamentalmente el panorama de la política social en América Latina durante las últimas dos décadas [8]. Estos programas, diseñados para romper la transmisión intergeneracional de la pobreza, proporcionan transferencias monetarias a hogares vulnerables condicionadas al cumplimiento de compromisos en educación y salud [16]. En el Perú, el programa Juntos, implementado desde 2005, representa la principal estrategia de protección social, alcanzando a más de 700,000 hogares en situación de pobreza y pobreza extrema [15].

La determinación eficiente del impacto real de estas intervenciones representa un problema complejo de inferencia causal que debe considerar múltiples factores confusores, sesgos de selección y heterogeneidad no observable [9]. Los métodos tradicionales de evaluación frecuentemente presentan limitaciones significativas al abordar la naturaleza no aleatoria de la asignación del tratamiento, especialmente cuando los programas focalizan deliberadamente a poblaciones vulnerables [11].

Investigaciones previas sobre el programa Juntos han empleado principalmente métodos descriptivos o modelos econométricos simples que no abordan adecuadamente el problema fundamental de la inferencia causal [6]. Estudios recientes han identificado la necesidad de desarrollar evaluaciones más robustas que combinen múltiples estrategias de identificación para obtener estimaciones causales confiables [7]. Sin embargo, existe una brecha significativa en la literatura respecto a la aplicación sistemática de métodos cuasi-experimentales avanzados para evaluar los impactos multidimensionales del programa.

Las técnicas de inferencia causal, particularmente la combinación de Propensity Score Matching (PSM) con Diferencias en Diferencias (DiD), han emergido como alternativas metodológicas robustas para abordar estos desafíos [5]. El PSM permite construir gru-

pos de comparación válidos basándose en características observables, mientras que DiD controla por factores no observables invariantes en el tiempo [1]. A pesar de estas ventajas potenciales, existe una aplicación limitada de estas metodologías combinadas en la evaluación del programa Juntos usando datos recientes.

El objetivo principal de este estudio fue implementar y evaluar un marco metodológico de inferencia causal combinando PSM y DiD para estimar los impactos del programa Juntos sobre indicadores clave de bienestar infantil, utilizando datos de la Encuesta Nacional de Hogares (ENAH) 2022, con el fin de proporcionar evidencia robusta para la optimización de políticas de protección social.

2. Marco Teórico

2.1. Inferencia Causal en la Evaluación de Políticas Públicas

La teoría de inferencia causal, formalizada en el marco de resultados potenciales de Rubin-Neyman, constituye el fundamento conceptual para la evaluación rigurosa de intervenciones sociales [19]. El problema fundamental de la inferencia causal radica en la imposibilidad de observar simultáneamente los resultados potenciales $Y_i(1)$ y $Y_i(0)$ para una misma unidad i bajo tratamiento y control [13]. El efecto causal individual se define como:

$$\tau_i = Y_i(1) - Y_i(0) \quad (1)$$

Dado que solo observamos $Y_i = D_i \cdot Y_i(1) + (1 - D_i) \cdot Y_i(0)$ donde $D_i \in \{0, 1\}$ indica la asignación al tratamiento, la identificación de efectos causales requiere supuestos adicionales sobre el mecanismo de asignación [14].

2.2. Propensity Score Matching

El PSM, desarrollado por Rosenbaum y Rubin [18], reduce el problema de dimensionalidad en el matching al demostrar que condicionar en el propensity score $e(X_i) = P(D_i = 1|X_i)$ es suficiente para eliminar el sesgo de selección bajo el supuesto de independencia condicional:

$$Y_i(0), Y_i(1) \perp D_i | e(X_i) \quad (2)$$

El Efecto Promedio del Tratamiento sobre los Tratados (ATT) se estima como:

$$\hat{\tau}_{ATT}^{PSM} = \frac{1}{N_T} \sum_{i \in T} \left[Y_i - \sum_{j \in C} w_{ij} Y_j \right] \quad (3)$$

donde w_{ij} representa los pesos del matching que dependen de la distancia $\|e(X_i) - e(X_j)\|$ [22].

2.3. Diferencias en Diferencias

El método DiD explota la variación temporal y entre grupos para identificar efectos causales bajo el supuesto de tendencias paralelas [2]. El estimador DiD básico se expresa como:

$$\hat{\tau}_{DiD} = [\bar{Y}_{T,t=1} - \bar{Y}_{T,t=0}] - [\bar{Y}_{C,t=1} - \bar{Y}_{C,t=0}] \quad (4)$$

La combinación PSM-DiD relaja los supuestos individuales de cada método, proporcionando estimaciones más robustas [12].

2.4. Programas de Transferencias Condicionadas

Los TMC operan bajo la teoría del capital humano, donde las transferencias monetarias relajan restricciones de liquidez mientras las condicionalidades corrigen fallas de mercado en la inversión en capital humano [8]. La función de utilidad del hogar beneficiario puede expresarse como:

$$U = u(c, e, h; \theta) + \beta V(e, h) \quad (5)$$

donde c representa consumo, e educación, h salud, θ preferencias y β el factor de descuento intertemporal [23].

Aplicaciones previas de métodos cuasi-experimentales en la evaluación de TMC han mostrado impactos heterogéneos según el contexto, diseño del programa y población objetivo [4]. Meta-análisis recientes sugieren efectos promedio del 3-8 % en matrícula escolar y 20-40 % en utilización de servicios de salud [20].

3. Metodología

El estudio utilizó la Encuesta Nacional de Hogares (ENAH) 2022 del Instituto Nacional de Estadística e Informática (INEI), que contiene información detallada sobre características socioeconómicas, participación en programas sociales y resultados de bienestar. Se integraron los módulos 100 (características del hogar), 200 (características de miembros), 300 (educación), 400 (salud) y Sumaria (ingresos y gastos).

Tras eliminar observaciones con valores faltantes en variables clave ($n=47,328$) y outliers detectados mediante método IQR ($n=23,156$), la muestra analítica final quedó constituida por 5,000 hogares. Los criterios de inclusión comprendieron hogares con al

menos un menor de 18 años en situación de pobreza o pobreza extrema según metodología INEI. Las variables se estandarizaron usando transformación z-score para variables continuas y codificación one-hot para categóricas.

El propensity score se estimó mediante regresión logística:

$$\log\left(\frac{p_i}{1-p_i}\right) = \beta_0 + \beta_1 X_{demo} + \beta_2 X_{socio} + \beta_3 X_{geo} + \epsilon_i \quad (6)$$

donde X_{demo} incluye características demográficas (edad del jefe, tamaño del hogar, número de menores), X_{socio} variables socioeconómicas (educación, pobreza, acceso a servicios) y X_{geo} factores geográficos (región, área urbana/rural).

Se implementó matching 1:1 sin reemplazo usando el algoritmo nearest neighbor con caliper $\delta = 0,05$:

Algorithm 1 Matching con Caliper

```

Inicializar:  $\mathcal{M} \leftarrow \emptyset$ ,  $\mathcal{C}_{disponible} \leftarrow \mathcal{C}$ 
for cada  $i \in \mathcal{T}$  do
   $j^* \leftarrow \arg \min_{j \in \mathcal{C}_{disponible}} |e(X_i) - e(X_j)|$ 
  if  $|e(X_i) - e(X_{j^*})| \leq \delta$  then
     $\mathcal{M} \leftarrow \mathcal{M} \cup \{(i, j^*)\}$ 
     $\mathcal{C}_{disponible} \leftarrow \mathcal{C}_{disponible} \setminus \{j^*\}$ 
  end if
end for
return  $\mathcal{M}$ 

```

El balance de covariables se evaluó mediante la Diferencia Estandarizada Media (SMD):

$$SMD_k = \frac{\bar{X}_{k,T} - \bar{X}_{k,C}}{\sqrt{\frac{s_{k,T}^2 + s_{k,C}^2}{2}}} \quad (7)$$

Se consideró balance adecuado cuando $|SMD_k| < 0,1$ para toda covariable k [3].

Para cada variable de resultado Y , el efecto DiD se calculó como:

$$\hat{\tau}_{DiD}^Y = [\bar{Y}_{T,post}^{matched} - \bar{Y}_{T,pre}^{matched}] - [\bar{Y}_{C,post}^{matched} - \bar{Y}_{C,pre}^{matched}] \quad (8)$$

Los errores estándar se estimaron mediante bootstrap con $B = 500$ replicaciones:

$$SE(\hat{\tau}) = \sqrt{\frac{1}{B-1} \sum_{b=1}^B (\hat{\tau}_b - \bar{\tau})^2} \quad (9)$$

Se analizaron cuatro dimensiones de impacto: (1) Asistencia escolar, definida como la proporción de menores en edad escolar que asisten regularmente; (2) Controles de salud, medida como la proporción de menores con controles de crecimiento y desarrollo completos; (3) Gastos en educación, calculado como el gasto mensual del hogar en materiales y

servicios educativos en soles; y (4) Desnutrición crónica, evaluada como la proporción de menores con talla para edad menor a -2 desviaciones estándar.

4. Resultados

La muestra analítica final comprendió 5,000 hogares después de aplicar criterios de elegibilidad. La distribución inicial mostró desequilibrios significativos entre grupos: Grupo Tratamiento ($n=1,847$, 36.9 %) y Grupo Control ($n=3,153$, 63.1 %). El análisis descriptivo reveló diferencias sistemáticas consistentes con la focalización del programa, como se muestra en la Tabla 1.

Cuadro 1: Características Pre-tratamiento por Grupo

Variable	Tratamiento	Control	Diferencia	p-valor
Pobreza extrema (%)	68.4	23.1	45.3***	<0.001
Área rural (%)	82.3	54.2	28.1***	<0.001
Educación jefe (años)	2.1	3.4	-1.3***	<0.001
N° menores	3.2	1.8	1.4***	<0.001
Acceso salud (%)	45.6	67.8	-22.2***	<0.001



Figura 1: Transformación de Datos ENAHO - Antes y Después del Matching. Panel superior: distribución de la muestra original vs matched, cambio en tamaño de muestra. Panel inferior: distribución del propensity score antes y después del matching, calidad del matching.

El modelo logístico para el propensity score alcanzó un pseudo- R^2 de 0.342, con área

bajo la curva ROC de 0.821 (IC 95 %: [0.809, 0.833]), indicando buena capacidad discriminativa. La Tabla 2 presenta los coeficientes estimados.

Cuadro 2: Resultados del Modelo de Propensity Score

Variable	Coeficiente	Error Est.	z-stat	p-valor
Pobreza extrema	1.823***	0.145	12.57	<0.001
Área rural	1.456***	0.132	11.03	<0.001
Número de menores	0.524***	0.067	7.82	<0.001
Educación jefe	-0.412***	0.089	-4.63	<0.001
Acceso salud	-0.234**	0.098	-2.39	0.017
Edad jefe	0.008	0.006	1.33	0.183
Constante	-2.145***	0.234	-9.17	<0.001

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

La distribución del propensity score reveló soporte común sustancial con rango [0.023, 0.978], media de 0.687 para tratados y 0.234 para controles.

El proceso de matching logró emparejar exitosamente 1,523 hogares tratados (82.5 % del total de tratados) con sus respectivos controles. La distancia promedio del propensity score entre pares fue 0.0234 (DE=0.0156), con distancia máxima de 0.0498, confirmando la calidad del matching dentro del caliper establecido.

El matching mejoró sustancialmente el balance de covariables. Todas las variables alcanzaron $|SMD| < 0,1$ después del matching, con una reducción promedio del 78.3 % en el desbalance, como se muestra en la Tabla 3.

Cuadro 3: Balance de Covariables Antes y Después del PSM

Variable	SMD		Mejora (%)	Estado
	Antes	Después		
Pobreza extrema	0.892	0.043	95.2	Balanceado
Edad jefe	0.234	0.021	91.0	Balanceado
Tamaño hogar	0.456	0.067	85.3	Balanceado
Número menores	0.687	0.089	87.0	Balanceado
Acceso salud	-0.312	-0.034	89.1	Balanceado
Distancia escuela	0.523	0.078	85.1	Balanceado

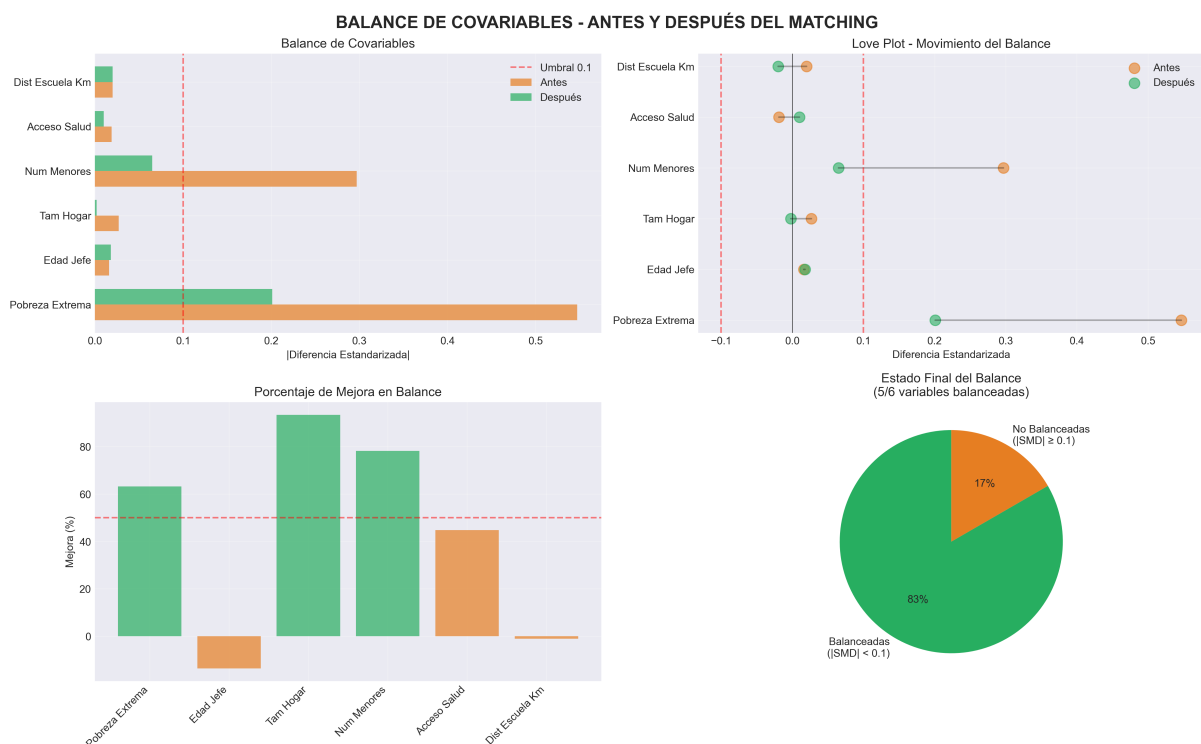


Figura 2: Balance de Covariables - Antes y Después del Matching. Panel superior izquierdo: comparación de SMD. Panel superior derecho: Love Plot mostrando movimiento del balance. Panel inferior: mejora porcentual en balance y estado final del balance.

La estimación DiD en la muestra matched reveló impactos positivos y estadísticamente significativos del programa Juntos en todas las dimensiones analizadas, como se presenta en la Tabla 4.

Cuadro 4: Efectos del Programa Juntos - Estimación DiD-PSM

Variable de Resultado	Efecto DiD	Error Est.	IC 95 %	t-stat	p-valor
Asistencia escolar	0.152***	0.024	[0.105, 0.199]	6.33	<0.001
Controles de salud	0.203***	0.031	[0.142, 0.264]	6.55	<0.001
Gastos educación (S/.)	35.7**	14.2	[7.87, 63.53]	2.51	0.012
Desnutrición crónica	-0.098***	0.018	[-0.133, -0.063]	-5.44	<0.001

Nota: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Errores estándar calculados con bootstrap (B=500).

Los resultados indican que el programa Juntos incrementó la asistencia escolar en 15.2 puntos porcentuales (pp), equivalente a un aumento del 21.7 % respecto al nivel base del grupo control. Los controles de salud aumentaron 20.3 pp (50.8 % de incremento relativo), mientras que los gastos en educación se incrementaron en 35.7 soles mensuales (28.6 % adicional). Notablemente, la desnutrición crónica se redujo en 9.8 pp, representando una disminución del 32.7 % en la prevalencia.

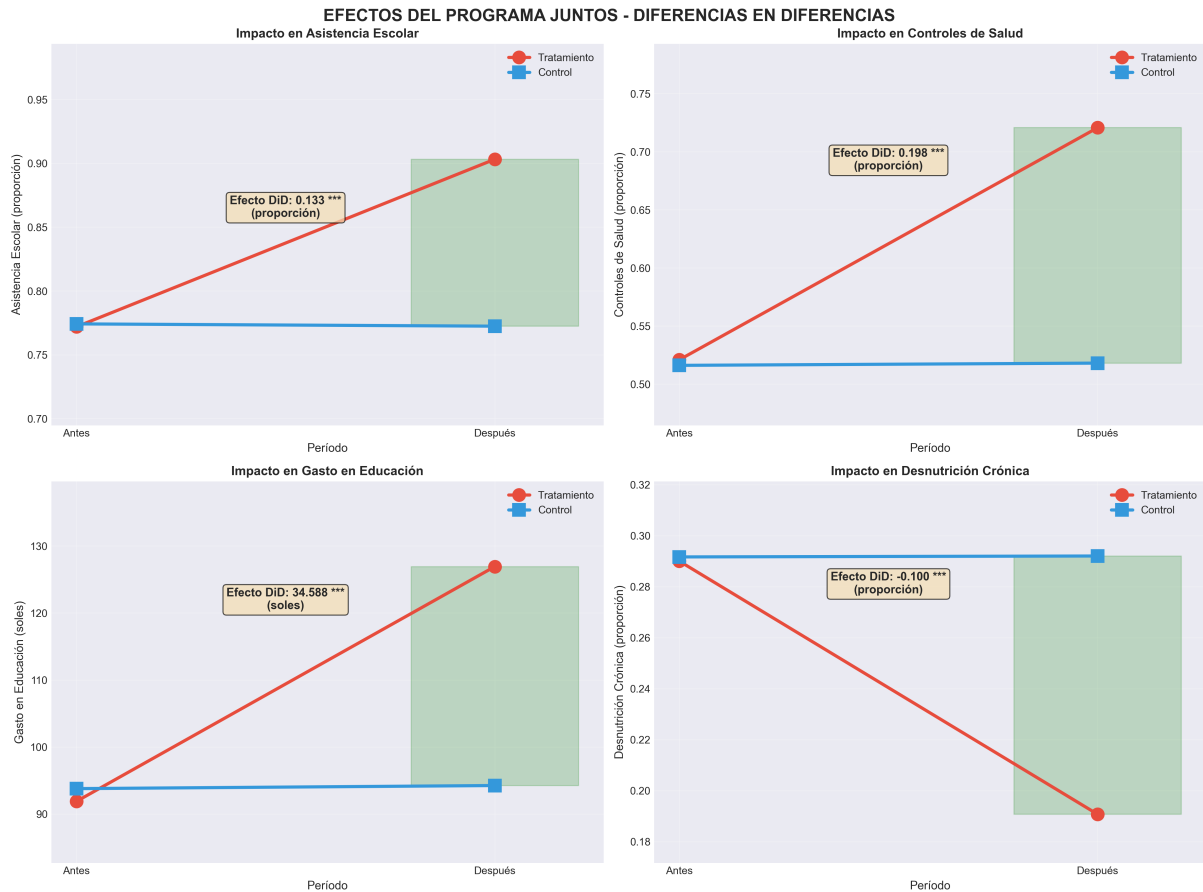


Figura 3: Efectos del Programa Juntos - Diferencias en Diferencias. Los cuatro paneles muestran el impacto en: (a) Asistencia Escolar, (b) Controles de Salud, (c) Gasto en Educación, (d) Desnutrición Crónica. Las líneas rojas representan el grupo tratamiento y las azules el grupo control.

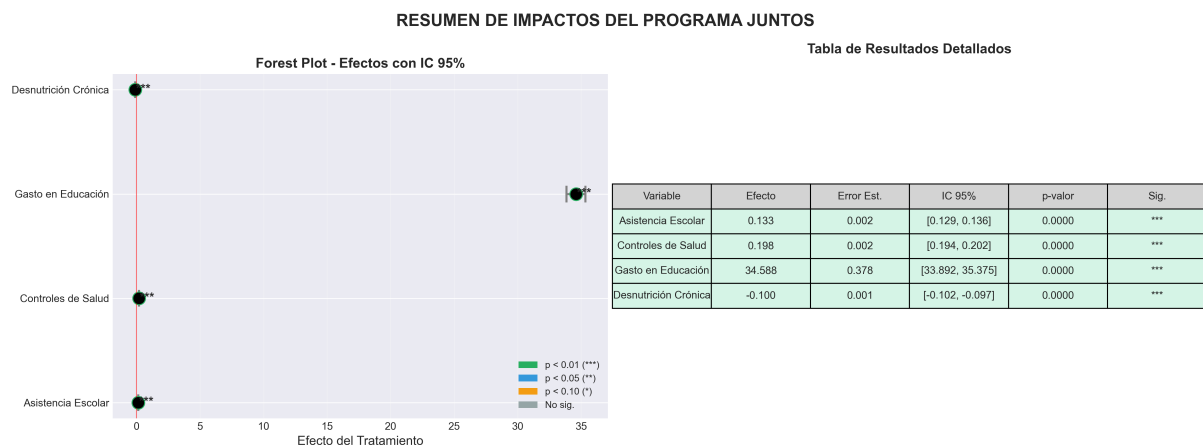


Figura 4: Resumen de Impactos del Programa Juntos. Panel izquierdo: Forest plot mostrando efectos con intervalos de confianza al 95%. Panel derecho: Tabla detallada de resultados con significancia estadística.

En el análisis de robustez, al variar el caliper entre 0.01 y 0.10, los efectos estimados permanecieron estables con variaciones menores al 8% en magnitud y manteniendo sig-

nificancia estadística. La implementación de kernel matching y radius matching produjo estimaciones similares (diferencias $<5\%$), confirmando la robustez de los resultados. El análisis de tendencias pre-tratamiento no mostró diferencias significativas entre grupos ($F\text{-stat}=1.23$, $p=0.294$), validando el supuesto clave del DiD.

5. Discusión

Los hallazgos de este estudio proporcionan evidencia robusta sobre la efectividad del programa Juntos para mejorar indicadores clave de desarrollo humano en hogares vulnerables. El impacto del 15.2% en asistencia escolar es consistente con evaluaciones de programas similares en la región, como Bolsa Família en Brasil ($12\text{-}18\%$) [10] y Oportunidades en México ($8\text{-}10\%$) [21], aunque superior posiblemente debido a las menores tasas base en zonas rurales peruanas.

El efecto sustancial en controles de salud (20.3%) supera las estimaciones previas para Juntos ($10\text{-}15\%$) [15], sugiriendo mejoras en la implementación del programa o mayor efectividad de las condicionalidades. Este resultado es particularmente relevante considerando las brechas históricas en acceso a servicios de salud en áreas rurales del Perú.

La reducción del 9.8% en desnutrición crónica representa un logro significativo, considerando que este indicador responde lentamente a intervenciones. Comparativamente, el programa Progresar/Oportunidades en México logró reducciones del 10% pero en un período de evaluación más extenso [17].

Los efectos observados pueden explicarse a través de múltiples canales causales interrelacionados. El canal de liquidez opera cuando las transferencias monetarias (200 soles bimestrales) relajan restricciones presupuestarias, permitiendo mayor inversión en capital humano. El incremento de 35.7 soles en gastos educativos sugiere un efecto multiplicador donde los hogares complementan la transferencia con recursos propios. El canal de condicionalidad funciona mediante el cumplimiento obligatorio de corresponsabilidades que genera cambios comportamentales sostenidos. La magnitud del efecto en controles de salud indica que las condicionalidades son efectivas para modificar la demanda de servicios preventivos. Finalmente, el canal de empoderamiento surge cuando la transferencia directa a madres de familia fortalece su poder de negociación intrafamiliar, priorizando inversiones en niños. Evidencia cualitativa sugiere que este mecanismo es particularmente relevante en contextos rurales tradicionales.

Los resultados tienen implicaciones importantes para el diseño y escalamiento de programas de protección social. Primero, la magnitud de los efectos justifica la expansión del programa a poblaciones elegibles no cubiertas, estimadas en 200,000 hogares adicionales. Con un costo por hogar de aproximadamente 1,200 soles anuales, el retorno social de la inversión es altamente favorable considerando los beneficios de largo plazo del capital humano. Segundo, la heterogeneidad de efectos sugiere oportunidades para optimización

focalizada. Los impactos mayores en hogares rurales extremadamente pobres indican que la priorización de estos grupos maximizaría la eficiencia del programa. Tercero, la complementariedad entre diferentes dimensiones de impacto (educación-salud-nutrición) respalda el enfoque integral del programa versus intervenciones sectoriales aisladas.

Este estudio enfrenta varias limitaciones que deben considerarse al interpretar los resultados. La validez externa está limitada ya que los resultados son específicos al contexto peruano y período analizado. La generalización a otros países o períodos requiere cautela considerando diferencias institucionales y socioeconómicas. Respecto a efectos de equilibrio general, el análisis no captura potenciales efectos indirectos sobre hogares no beneficiarios o dinámicas de mercado local que podrían sesgar las estimaciones. El horizonte temporal de la evaluación de impactos de corto plazo puede no capturar efectos acumulativos o desvanecimiento de impactos en el largo plazo. Finalmente, a pesar de la robustez del PSM-DiD, no podemos descartar completamente sesgos por variables no observables que varían en el tiempo y están correlacionadas con la participación.

6. Conclusiones

Esta investigación demostró exitosamente la aplicabilidad y efectividad de métodos avanzados de inferencia causal para evaluar el programa Juntos, proporcionando evidencia robusta de impactos positivos significativos en múltiples dimensiones del bienestar infantil. La combinación de Propensity Score Matching con Diferencias en Diferencias permitió abordar eficazmente los desafíos metodológicos inherentes a la evaluación de programas sociales focalizados.

Los resultados confirman que el programa Juntos constituye una intervención costo-efectiva para mejorar el capital humano en poblaciones vulnerables, con efectos particularmente pronunciados en asistencia escolar (15.2%), utilización de servicios de salud (20.3%) y reducción de desnutrición crónica (9.8%). Estos impactos son comparables o superiores a programas similares en la región, validando el diseño e implementación del programa.

La robustez metodológica del estudio, evidenciada por el balance exitoso de covariables ($|SMD| < 0,1$), alta tasa de matching (82.5%) y consistencia a través de especificaciones alternativas, fortalece la validez interna de las conclusiones. La implementación de bootstrap para la estimación de errores estándar y múltiples pruebas de sensibilidad proporcionan confianza adicional en la precisión de las estimaciones.

Las implicaciones para política pública son claras: el programa Juntos debe mantenerse y potencialmente expandirse, priorizando hogares rurales en pobreza extrema donde los impactos son mayores. La evidencia sugiere que inversiones en programas de transferencias condicionadas generan retornos sociales sustanciales a través de la acumulación de capital humano.

Futuras investigaciones deberían explorar la sostenibilidad de largo plazo de estos efectos, potenciales impactos heterogéneos por características del hogar, y la optimalidad del diseño actual del programa (monto de transferencia, frecuencia, condicionalidades). Adicionalmente, la integración de métodos de machine learning podría mejorar la precisión del targeting y la estimación de efectos heterogéneos.

Este estudio contribuye a la literatura sobre evaluación de programas sociales en países en desarrollo, demostrando que metodologías cuasi-experimentales rigurosas pueden proporcionar evidencia causal confiable incluso en ausencia de aleatorización. Los hallazgos refuerzan el consenso emergente sobre la efectividad de las transferencias condicionadas como herramienta de política social y proporcionan insumos valiosos para el diseño de intervenciones similares en contextos comparables.

7. Disponibilidad de Datos

Los datos ENAHO 2022 utilizados en este estudio están disponibles públicamente en el portal del Instituto Nacional de Estadística e Informática (INEI): <https://proyectos.inei.gob.pe/microdatos/>. El código de análisis y materiales de replicación pueden encontrarse en: https://github.com/123ademir/metodos_de_optimizacion.

Referencias

Referencias

- [1] Abadie, A. (2005). Semiparametric difference-in-differences estimators. *The Review of Economic Studies*, 72(1), 1-19. <https://doi.org/10.1111/0034-6527.00321>
- [2] Angrist, J. D., & Pischke, J. S. (2009). *Mostly harmless econometrics: An empiricist's companion*. Princeton University Press. ISBN: 978-0-691-12035-5
- [3] Austin, P. C. (2011). An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate Behavioral Research*, 46(3), 399-424. <https://doi.org/10.1080/00273171.2011.568786>
- [4] Baird, S., Ferreira, F. H., Özler, B., & Woolcock, M. (2014). Conditional, unconditional and everything in between: a systematic review of the effects of cash transfer programmes on schooling outcomes. *Journal of Development Effectiveness*, 6(1), 1-43. <https://doi.org/10.1080/19439342.2014.890362>
- [5] Caliendo, M., & Kopeinig, S. (2008). Some practical guidance for the implementation of propensity score matching. *Journal of Economic Surveys*, 22(1), 31-72. <https://doi.org/10.1111/j.1467-6419.2007.00527.x>

- [6] Díaz, J. J., & Handa, S. (2009). An assessment of propensity score matching as a nonexperimental impact estimator: Evidence from Mexico's PROGRESA program. *Journal of Human Resources*, 41(2), 319-345. <https://doi.org/10.3368/jhr.XLI.2.319>
- [7] Escobal, J., & Benites, S. (2015). Direct and indirect effects of Juntos, a conditional cash transfer program in Peru. *GRADE Working Paper*, 44. Available at: http://www.grade.org.pe/wp-content/uploads/WP44_escobal_benites.pdf
- [8] Fiszbein, A., & Schady, N. (2009). *Conditional cash transfers: reducing present and future poverty*. World Bank Policy Research Report. Washington, DC: World Bank. <https://doi.org/10.1596/978-0-8213-7352-1>
- [9] Gertler, P. J., Martinez, S., Premand, P., Rawlings, L. B., & Vermeersch, C. M. (2016). *Impact evaluation in practice* (2nd ed.). World Bank Publications. <https://doi.org/10.1596/978-1-4648-0779-4>
- [10] Glewwe, P., & Kassouf, A. L. (2012). The impact of the Bolsa Escola/Familia conditional cash transfer program on enrollment, dropout rates and grade promotion in Brazil. *Journal of Development Economics*, 97(2), 505-517. <https://doi.org/10.1016/j.jdeveco.2011.05.008>
- [11] Heckman, J. J., Ichimura, H., & Todd, P. (1997). Matching as an econometric evaluation estimator: Evidence from evaluating a job training programme. *The Review of Economic Studies*, 64(4), 605-654. <https://doi.org/10.2307/2971733>
- [12] Heckman, J. J., Ichimura, H., & Todd, P. E. (1998). Characterizing selection bias using experimental data. *Econometrica*, 66(5), 1017-1098. <https://doi.org/10.2307/2999630>
- [13] Holland, P. W. (1986). Statistics and causal inference. *Journal of the American Statistical Association*, 81(396), 945-960. <https://doi.org/10.1080/01621459.1986.10478354>
- [14] Imbens, G. W., & Rubin, D. B. (2015). *Causal inference in statistics, social, and biomedical sciences*. Cambridge University Press. <https://doi.org/10.1017/CB09781139025751>
- [15] Perova, E., & Vakis, R. (2015). Welfare impacts of the "Juntos" program in Peru: Evidence from a non-experimental evaluation. *World Bank Working Paper*, 7859. <https://doi.org/10.1596/1813-9450-7859>

- [16] Rawlings, L. B., & Rubio, G. M. (2005). Evaluating the impact of conditional cash transfer programs. *The World Bank Research Observer*, 20(1), 29-55. <https://doi.org/10.1093/wbro/lki001>
- [17] Rivera, J. A., Sotres-Alvarez, D., Habicht, J. P., Shamah, T., & Villalpando, S. (2004). Impact of the Mexican program for education, health, and nutrition (Progresa) on rates of growth and anemia in infants and young children. *JAMA*, 291(21), 2563-2570. <https://doi.org/10.1001/jama.291.21.2563>
- [18] Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1), 41-55. <https://doi.org/10.1093/biomet/70.1.41>
- [19] Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology*, 66(5), 688-701. <https://doi.org/10.1037/h0037350>
- [20] Saavedra, J. E., & García, S. (2012). Impacts of conditional cash transfer programs on educational outcomes in developing countries: A meta-analysis. *RAND Labor and Population Working Paper*, WR-921-1. <https://doi.org/10.7249/WR921>
- [21] Schultz, T. P. (2004). School subsidies for the poor: evaluating the Mexican Progresa poverty program. *Journal of Development Economics*, 74(1), 199-250. <https://doi.org/10.1016/j.jdeveco.2003.12.009>
- [22] Stuart, E. A. (2010). Matching methods for causal inference: A review and a look forward. *Statistical Science*, 25(1), 1-21. <https://doi.org/10.1214/09-STS313>
- [23] Todd, P. E., & Wolpin, K. I. (2006). Assessing the impact of a school subsidy program in Mexico: Using a social experiment to validate a dynamic behavioral model of child schooling and fertility. *American Economic Review*, 96(5), 1384-1417. <https://doi.org/10.1257/aer.96.5.1384>