

La etnicidad en el Perú y su naturaleza multidimensional: una propuesta de medición

Ethnicity in Peru and its multidimensional nature: a proposal for measurement

Guiliana Rocío Campos Briceño¹

Universidad Nacional Mayor de San Marcos. Lima, Perú

guiliana.campos@unmsm.edu.pe

ORCID: 0000-0002-1818-7358

Ronny Martin Condor Iturrizaga²

Universidad Nacional Mayor de San Marcos. Lima, Perú

ronny.condor@unmsm.edu.pe

ORCID: 0000-0002-2299-4566

Citar como: Campos, G. y Condor, R. (2022). La etnicidad en el Perú y su naturaleza multidimensional: una propuesta de medición. *Desde el Sur*, 14(1), e00012

RESUMEN

La etnicidad es un concepto complejo de definir en las ciencias sociales. Esta complejidad se debe a que diversos autores reconocen a la etnicidad como un conjunto de características raciales, sociales y culturales que comparten un grupo de personas. Debido a ello, la naturaleza misma de la etnicidad es multidimensional. El presente documento tiene como objetivo construir un índice que mida la multidimensionalidad de la etnicidad usando datos de la encuesta de hogares de Perú en 2007 y 2017. El método usado para construcción del índice es el análisis de componentes principales (ACP). Además, se presenta un análisis correlacional del índice propuesto con la incidencia de la pobreza, variables educativas y laborales. Entre los resultados encontrados, tenemos que hay una mayor

1 Bachiller en Economía por la Universidad Nacional Mayor de San Marcos (UNMSM). En 2020 se incorporó a Innovations for Poverty Action (IPA), donde se desempeña como asociada de investigación y ha dirigido proyectos relacionados con educación e inclusión financiera.

2 Bachiller en Economía por la UNMSM. Se desempeña como asociado de investigación en Innovations for Poverty Action (IPA). Ha sido consultor de la Contraloría General de la República y jefe de prácticas de los cursos de Econometría y Economía de la Distribución y Desigualdad en la UNMSM.

intensidad étnica en la población en situación de pobreza, en condición de analfabetismo y con menores niveles educativos, así como en la población autoempleada.

PALABRAS CLAVE

Etnicidad, grupos étnicos, multidimensional, índice

ABSTRACT

Ethnicity is a complex concept to define in the social sciences. This complexity is since various authors recognize ethnicity as a set of racial, social, and cultural characteristics shared by a group of people. As a result, the nature of ethnicity is multidimensional. This paper aims to construct an index that measures the multidimensionality of ethnicity using household survey data from Peru in 2007 and 2017. The method used to construct the index is Principal Component Analysis (PCA). In addition, a correlational analysis of the proposed index with the incidence of poverty, educational and labor variables is presented. Among the results found, there is a greater ethnic intensity in the population living in poverty, in illiteracy conditions and with lower educational levels, as well as in the self-employed population.

KEYWORDS

Ethnicity, ethnic groups, multidimensionality, index

Introducción

El Perú es reconocido como un país multiétnico y pluricultural. Se hablan 48 idiomas distintos y se cuentan con 7250 localidades de pueblos originarios³, según la Base de Datos de Pueblos Indígenas realizado por el Ministerio de Cultura. Además, dependiendo del indicador de etnicidad, Sulmont (2010) ha estimado que entre el 19,2 % y el 74,8 % de la población es indígena.

La gran diversidad de culturas está asociada a las grandes diferencias en los ingresos y oportunidades económicas (Pasquier-Doumer y Risso Brandon, 2015). La incidencia de pobreza en etnias según lengua materna es mayor por 21,6 puntos porcentuales respecto a los que solo hablan

3 Corresponde a los espacios geográficos donde habitan o ejercen sus derechos colectivos el o los pueblos indígenas u originarios, sea en propiedad o debido a otros derechos reconocidos por el Estado o que usan u ocupan tradicionalmente.

castellano (Presidencia del Consejo de Ministros, 2013). Respecto a la diferencia de oportunidades en el mercado laboral, un estudio realizado por Galarza y Yamada (2012) encontró evidencia de discriminación según apariencia física hacia los candidatos con apellidos indígenas y fue más notorio en carreras profesionales.

Se requiere investigaciones que aborden el estudio de la realidad de los grupos sociales. Si bien el concepto de etnicidad puede resultar complejo, y las definiciones pueden variar dependiendo del autor, es cierto que se tiende a considerar que la etnicidad está compuesta por múltiples dimensiones (Burton *et al.*, 2010; Kolo, 2012). Al hablar de la multidimensionalidad en la etnicidad, nos referimos a las numerosas características o expresiones que son compartidas por un mismo grupo. Podríamos mencionar algunas dimensiones de la etnicidad, como la raza, el color de piel, la familia lingüística, la ciudadanía y las costumbres. Reconocer la multidimensionalidad de la etnicidad y su importancia en el debate nos permitirá estudiar a los grupos étnicos, no de una manera fija y única, sino bajo una mirada más global.

La creciente literatura empírica sobre la etnicidad se ha centrado en tomar como insumos para su medición a las características raciales percibidas por terceros, autopercepción de raza, idioma y las tradiciones culturales. Estas dimensiones de identificación para definir colectivos étnicos son tomados para la cuantificación empírica (Burton *et al.*, 2010; Kolo, 2012). Siguiendo a Schkolnik (2009), se debe tener como mínimo cuatro dimensiones para una buena medición étnica.

La principal pregunta de investigación que busca responder este estudio es si existe la posibilidad de medir la etnicidad de tal manera que considere su naturaleza multidimensional. Además, el estudio busca responder cómo esta medición de etnicidad se relaciona con otras variables educativas y laborales, comparando los resultados de la medición propuesta con las medidas usadas en estudios previos. Por consiguiente, el presente documento tiene como principal objetivo contribuir en la comprensión y medición de las múltiples dimensiones de la etnicidad y resumirlas en un solo indicador, que llamaremos índice de etnicidad.

La importancia del estudio se centra en medir y capturar la multidimensionalidad de la etnicidad en un índice. Reconociendo que se puede cuantificar la etnicidad, se realiza un análisis correlacional del índice de etnicidad con la condición de pobreza, el nivel educativo y los años de educación alcanzados, la condición laboral y el salario mensual. El análisis mostrado deberá ser interpretado teniendo en cuenta que solo se trata de un análisis correlacional.

El índice de etnicidad tendrá cuatro componentes en su construcción y la metodología a usar será a través de un análisis de componentes principales (ACP). El índice de etnicidad será interpretado como la intensidad en las características —o dimensiones— relacionadas con la etnicidad. Un mayor valor del índice se interpreta como una mayor intensidad en sus características étnicas y, por tanto, se relacionará con un «mayor» indigenismo del individuo. Las personas con valores más bajos del índice tendrán, en cambio, menor intensidad en dimensiones asociadas a ser indígena.

Se encontró que de 2007 a 2017 el índice a nivel nacional ha disminuido, lo que da a entender que existiría una menor diversidad étnica. Asimismo, el índice presenta una mayor intensidad en conjuntos geográficos como sierra sur o sierra centro; en específico, en los departamentos de Huancavelica, Apurímac y Ayacucho.

El artículo sigue la siguiente estructura: la sección 2 explica la etnicidad en el Perú a través de algunos indicadores descriptivos y literatura asociada. De la misma forma, se introduce el concepto de índice de etnicidad y la metodología que sigue para su construcción. En la sección 3 se presentan algunos resultados que interrelacionan el índice con el logro educativo, el mercado laboral y el nivel de pobreza.

Etnicidad en el Perú

Los procesos y las dimensiones de la etnicidad en el Perú son aspectos que han sido largamente estudiados por diversos autores. Los grupos étnicos se basan por la geografía e historia de los individuos (Figueroa y Barrón, 2005). De igual forma, el lenguaje y el color de piel resultan relevantes como factores de análisis (Paredes y House, 2007). Entonces, teniendo en cuenta que las características individuales, tanto observables como inobservables, definen a los grupos étnicos, a continuación, se explicarán algunos estadísticos descriptivos para el caso peruano.

Indicadores de etnicidad

La etnicidad se puede medir por autoidentificación, es decir, a través de las percepciones sobre cómo se ven las personas en sí tal como una medida de raza (Roth, 2016). Otra forma de medición se realiza mediante indicadores de las características observables. Los datos que hacen posible este ejercicio de medición los encontramos en la Encuesta Nacional de Hogares (ENAH) realizado por el Instituto Nacional de Estadística (INEI). Por ello, se elaboró la tabla 1, donde se puede ver que en 2017 un 47,2 %

de la población se identificó como mestizo⁴, seguido por un 25,1 % que se considera quechua. Además, de 2007 a 2017 se aprecia que el porcentaje que se considera quechua va disminuyendo. En el mismo periodo, el porcentaje de las personas que se consideran mestizas, en promedio, es del 52,2 %.

En línea a Figueroa y Barrón (2005), la autoidentificación puede resultar contradictoria, pues en una sociedad donde el grupo dominante tienden a ser los blancos, es posible que las personas oculten ser *no blancos*. Saavedra *et al.* (2004) indican que la autoidentificación mestiza de los diferentes grupos étnicos genera que esta categoría se vuelva heterogénea. Por ello, los autores construyeron un puntaje en el que comprueban que los individuos se declaran más blancos de lo que realmente son, según puntuaciones de sus características fenotípicas. En este sentido, es importante complementar el análisis de la etnicidad del Perú con el estudio de otras variables.

TABLA 1. Grupos étnicos según autoidentificación (en porcentaje)

Año	Quechua	Aimara	Nativo de la Amazonía	Negro	Blanco	Mestizo	Otro	No sabe / No responde
2007	31,0 %	3,5 %	2,1 %	2,0 %	4,0 %	54,7 %	2,9 %	0,0 %
2008	31,2 %	3,4 %	2,2 %	1,4 %	4,2 %	52,5 %	5,1 %	0,0 %
2009	30,2 %	3,2 %	2,2 %	2,2 %	4,5 %	53,9 %	3,9 %	0,0 %
2010	29,7 %	3,3 %	2,2 %	2,1 %	4,3 %	52,9 %	3,8 %	1,8 %
2011	29,4 %	3,1 %	1,9 %	1,9 %	3,6 %	52,9 %	2,6 %	4,7 %
2012	26,8 %	3,0 %	1,7 %	1,8 %	4,8 %	52,2 %	6,0 %	3,8 %
2013	25,0 %	2,9 %	1,5 %	1,6 %	4,3 %	52,8 %	6,9 %	5,0 %
2014	23,7 %	2,9 %	1,6 %	2,1 %	4,6 %	52,4 %	6,7 %	6,0 %
2015	23,8 %	2,7 %	1,7 %	2,2 %	4,5 %	52,5 %	5,7 %	7,0 %
2016	23,5 %	3,0 %	1,8 %	2,4 %	4,5 %	51,4 %	6,6 %	6,9 %
2017	25,1 %	2,8 %	1,7 %	6,3 %	4,4 %	47,2 %	5,9 %	6,5 %
Total	26,9 %	3,0 %	1,8 %	2,4 %	4,4 %	52,2 %	5,2 %	4,0 %

Fuente: Encuesta Nacional de Hogares
Elaboración propia

Según datos de la misma encuesta, en 2017 un 17,8 % indicó que habla otra familia lingüística que no sea el español. Es interesante ver que, de esta última cifra, un 33,2 % es pobre extremo. En cuanto al nivel educativo,

4 Según el INEI (2017), mestizo se define como la combinación de blanco con indígena.

el 24,2 % no cuenta con nivel educativo y el 25,4 %, solo tiene nivel primaria. Por el lado laboral, un 33,7 % declaró ser trabajador familiar no remunerado.

En la tabla 2, siguiendo la definición de De Tas, Reimao y Orlando (2014), se define como indígenas a las personas que declaran tener como lengua materna al quechua, aimara o alguna lengua amazónica. Según datos del INEI, en 2017 la tasa de pobreza es menor en el grupo no indígena por 14 puntos porcentuales. Asimismo, se puede indicar que el nivel educativo, o años de educación, va a ser mayor en el grupo no indígena. Sobre la condición laboral, un 48,3 % del grupo étnico indica ser autoempleado en su centro de trabajo, en decir, hay una mayor proporción de la población no indígena en el sector salarial. Estos datos reflejan que existen brechas educativas, de ingresos y oportunidades económicas en los grupos multiétnicos para el Perú y que estas brechas, incluso, pueden ser mayores por área de residencia (Pasquier-Doumer y Risso Brandon, 2015; Saavedra *et al.*, 2004).

TABLA 2. Distribución de la población, pobreza, mercado laboral y educación según grupos étnicos

	Distribución	Nivel de pobreza	Condición laboral	Educación	
			Autoempleado	Tasa de analfabetismo	Nivel de educación superior a la secundaria completa
Indígena	17 %	37 %	48 %	20 %	27 %
Urbano	56 %	22 %	47 %	15 %	36 %
Rural	44 %	47 %	49 %	28 %	15 %
No indígena	83 %	19 %	33 %	4 %	47 %
Urbano	85 %	13 %	32 %	3 %	52 %
Rural	15 %	39 %	42 %	14 %	16 %
Total	100				

Fuente: Encuesta Nacional de Hogares (2017)

Elaboración propia

Como se puede ver en la tabla 2, las diferencias en cuanto a condición laboral y educación resultan reveladoras cuando se analizan desde la perspectiva de la etnicidad. Por ello, es importante realizar una medida —o índice— para tratar de aproximarnos a las brechas sociales que se pueden dar entre los grupos étnicos.

Desigualdad étnica

La literatura en cuanto a desigualdad étnica puede estudiar: i) la relación entre la diversidad étnica y el desarrollo (Alesina *et al.*, 2016) y ii) sus efectos en las brechas de logro educativo, ingresos y oportunidades. Los resultados sociales de esta desigualdad pueden reflejarse en la discriminación y en la exclusión de los individuos (Galarza y Yamada Fukusaki, 2012).

Sobre la desigualdad étnica en grupos racializados, Woo-Mora (2021) encuentra que las inequidades entre grupos racializados tienen un afecto adverso sobre el bienestar de la sociedad. Un interesante aporte de los autores es que el incremento de 1 % en la proporción de la desigualdad en ingresos de grupos racializados sobre el total de ingresos se relaciona con la reducción del PBI per cápita en un 4 %.

La desigualdad étnica puede medirse por un índice que pondere las diferencias en los ingresos de los grupos étnicos. Siguiendo ello, Ravallion (2019) ha encontrado que la disminución de la desigualdad ha contribuido con reducción de la pobreza en Malasia en los últimos 50 años. A pesar de que no ha sido tan significativo desde el punto de vista cuantitativo para la reducción de la pobreza como lo ha sido el crecimiento general, eso no implica que la redistribución étnica no sea una herramienta contundente contra la pobreza.

Telles *et al.* (2015) encuentran evidencia de que en ocho países latinoamericanos, entre ellos el Perú, la autoidentificación étnica de grupos indígenas o negros se relaciona negativamente con la escolarización, aunque no siempre es estadísticamente significativa por país.

Metodología

Al hablar de etnicidad, no nos referimos solamente a una característica racial. La raza hace referencia a características observables de una persona (como su color de piel, por ejemplo), mientras que la etnicidad se refiere a cultura y, específicamente, a diferencias culturales (Bello y Rangel, 2000). La cultura, además de expresarse en características observables, es también expresada en características inobservables, como el modo de comprender el mundo, el sentido de pertenencia a un grupo y las creencias, que influyen en la forma de cómo un grupo social con similares características comparte ciertas tradiciones, costumbres y modos de vida. Por tanto, al hablar de etnicidad, debemos considerar la multidimensionalidad como parte de su definición, lo cual implicaría que, además de la raza, las características étnicas interactúan también con la cultura, la religión, el lenguaje, las tradiciones y otras dimensiones más (Saavedra *et al.*, 2004).

En otro contexto, la religión sería considerada también como parte de la construcción del índice de etnicidad; sin embargo, el Perú es un país mayoritariamente católico, incluso en grupos étnicos distintos (Figueroa y Barrón, 2005), y, por lo tanto, no sería una variable que genere distinción entre un grupo y otro.

Una característica importante para comprender la etnicidad es la lengua materna, la cual, dependiendo del lugar donde se creció, puede ser distinta. Sin embargo, debido a que en Perú el idioma predominante es el castellano⁵, y, además, las familias lingüísticas nativas están desapareciendo cada vez más, esta variable es una medida parcial de etnicidad.

Asimismo, las personas, por sus raíces y costumbres, pueden autoidentificarse con algún grupo étnico. Es importante reconocer la limitación de esta medida y los sesgos que puede tener. Ñopo *et al.* (2007) desarrollan el concepto de intensidad racial y muestran que existen diferencias significativas entre la intensidad de la raza registrada por una persona externa (un encuestador) y la que es autorregistrada. Es decir, existían personas que se identificaban más blancas o menos indígenas de lo que registraba el encuestador⁶, lo cual refuerza el argumento de la limitación de esta variable como una medida completa de etnicidad. Además, González Huamán *et al.* (2021), al estudiar el rol del empadronador del censo de 2017 en Perú, encuentra que una gran parte de los empadronadores no estaban debidamente entrenados para aplicar la pregunta sobre autoidentificación étnica. La mayoría explicó esta pregunta a los entrevistados a su modo de entenderla, basándose en su conocimiento y experiencia previa, mientras que otros reformularon esta pregunta con la intención de hacerla más comprensible para el entrevistado, lo cual conduciría a un sesgo en la medición de la etnicidad. Aunque la autoidentificación de los individuos puede diferenciarse de la raza observada por una persona externa, ambas son medidas de raza y, tal como enfatiza Roth (2016), no existe una «correcta» o «verdadera» medida, lo importante es conocer el límite de la variable a tomar para la investigación posterior que se realice con ella. La autora también comenta que el estudio de la raza nos permite entender las oportunidades desiguales y las interacciones complejas entre los colectivos de personas.

5 Según datos de la ENAHO de 2019, el 79,1 % reportó que el castellano es su lengua materna. Este porcentaje, en cambio, fue de 77,6 % en 2004.

6 El estudio de Ñopo complementa la autoidentificación con la identificación externa por parte de un encuestador adecuadamente capacitado para puntuar al individuo según su intensidad racial. Esto implicó un trabajo de campo para recoger esta información sobre intensidad racial. La ENAHO no contiene un módulo similar.

Para Burton *et al.* (2010), el país de nacimiento, o la condición de migrante, es también una dimensión a tomar en cuenta cuando hablamos de etnicidad. Al reconocer las diferencias existentes entre los seres humanos, somos conscientes también de que estas se pueden acrecentar si es que una persona convive en un país distinto, con costumbres que le resultan extrañas, lo que conllevaría a que la persona no se sienta comprendida y no encaje en la nueva sociedad de la que es parte (Geertz, 1993). El migrante, al verse expuesto a distintas culturas, puede cambiar su identidad étnica, lo que reforzaría el argumento de que la etnicidad es un concepto dinámico (Phinney, 1990). Además, los migrantes y las minorías tienden a recrear instituciones y estructuras que les permitan conservar y reproducir sus tradiciones, cultura y valores religiosos (Ahmad y Bradby, 2007). El país de origen juega un rol importante en la identificación étnica del migrante, pues muchos tienden agruparse con sus similares y reconocer al país de origen, su nación, como su grupo étnico (Martin y Gerber, 2006; McDonnell y De Lourenco, 2009). En la presente investigación, usamos la condición de migrante departamental, como una dimensión de la etnicidad. Aquellas personas que migran de un departamento a otro estarán expuestas a una diversidad de tradiciones y culturas, lo que le permitiría identificarse con distintos grupos étnicos. Siguiendo a Giraldo *et al.* (2019), en esa delgada línea entre el espacio de origen y el espacio de destino, en los migrantes internos y sus familias, ocurre un proceso de transculturización entre el origen provinciano-rural y la identificación capitalina-urbana.

De acuerdo con el INEI, un área rural es un territorio integrado por centros poblados rurales, asentamientos y localidades rurales, que no tiene más de 100 viviendas agrupadas contiguamente ni es capital de distrito; o que, teniendo más de 100 viviendas, estas se encuentran dispersas o diseminadas sin formar bloques o núcleos. Por lo general, las áreas rurales no están completamente desarrolladas y, a pesar de la notable reducción de la pobreza a partir de 2000, esto se concentró en áreas urbanas, mientras que el 45 % de la población rural sigue siendo pobre (Morley, 2017, p. 2). La población rural, además, tienen tradiciones lingüísticas y culturales propias que afectan su relación con la población nacional de la que forman parte (Trivelli *et al.*, 2000, p. 20). A pesar de que en ciertas áreas rurales puede estar presente la predominante autoidentificación mestiza (Sulmont, 2011), para la construcción del índice, consideramos la ruralidad como una dimensión, asumiendo que en estas áreas la estructura y organización son tradicionales, además de que se concentran aquí mayores expresiones culturales ancestrales.

Reconociendo la multidimensionalidad de la etnicidad, proponemos una medida que considere distintas características y las agrupe en una

sola medida, que en adelante llamaremos «índice de etnicidad». La principal fuente de información es la ENAHO del INEI. La ENAHO es aplicada desde 1997; sin embargo, a partir de 2004 se usa una metodología actualizada, que es la que usaremos para el presente estudio. El periodo analizado será de 2007 a 2017. La ENAHO cuenta con diversos módulos que buscan capturar información sobre las condiciones de vida; tal es así, que usaremos las preguntas relacionadas con etnicidad para construir nuestro índice. En la tabla 3 podemos observar las variables que componen el índice propuesto. Además, otros módulos de esta encuesta nos permitirán asociar el índice de etnicidad con otras variables, como condición de pobreza, variables educativas y del mercado de trabajo.

TABLA 3. Composición de índice de etnicidad

Variable	Descripción
Grupo étnico autoidentificado	Por sus antepasados y por sus costumbres, usted se considera: 0 = No indígena • Blanco • Mestizo 1 = Indígena • Quechua • Aimara • Nativo de la Amazonía • Negro/moreno/zambo/mulato/afroperuano
Lengua materna	Lengua materna que aprendió en su niñez: 0 = No indígena • Castellano • Inglés • Portugués • Otra lengua extranjera 1 = Indígena • Quechua • Aimara • Otro dialecto nativo
Migración departamental	Migró a un departamento distinto al que nació: 0 = No 1 = Sí
Ruralidad	Estrato geográfico: 0 = Urbano 1 = Rural

La metodología usada para construir el índice de etnicidad es el ACP, una técnica que busca formar nuevas variables que son combinaciones

lineales de las variables iniciales (Sharma, 1996). Una representación formal del ACP para n variables es la siguiente:

$$\begin{aligned} PC_i &= a_{i1}X_1 + a_{i2}X_2 + \dots + a_{in}X_n \\ &\vdots \\ PC_k &= a_{k1}X_1 + a_{k2}X_2 + \dots + a_{kn}X_n \end{aligned}$$

Donde a_{kn} representa el peso para el componente principal k y la variable n .

El primer componente explicaría la mayor varianza de las variables iniciales, sujeta a que la suma cuadrada de los pesos sea igual a 1; el segundo componente explicaría la segunda mayor varianza y así sucesivamente. Los pesos y la varianza de cada componente principal se determinan con base en la matriz de correlación⁷ de las variables iniciales (Vyas y Kumara-nayake, 2006).

El ACP es una técnica ampliamente usada en estudios que buscan generar índices de nivel socioeconómico. En este estudio, en cambio, aplicamos esta metodología convenientemente para construir el índice de etnicidad. Las nuevas variables creadas a partir del ACP, también llamadas «componentes», tienen la principal característica que no están correlacionadas entre sí, lo cual evitaría el problema de multicolinealidad⁸ en un análisis de regresión. Además, el ACP permitirá identificar patrones de asociación que, inicialmente, no sean distinguibles al analizar las variables iniciales (Martel *et al.*, 2021). En la tabla 4 podemos observar la matriz de los coeficientes de correlación entre las variables que componen nuestro índice. Podemos ver también que los coeficientes de correlación son significativos a un nivel del 1 %. Esto muestra la existencia de una fuerte correlación entre las variables étnicas que identificamos y, como se mencionó, podría provocar problemas de multicolinealidad en un análisis de regresión.

7 Si bien en esta investigación no hemos desarrollado la parte formal del ACP, una explicación más detallada sobre la «eigenestructura» de la matriz de correlación de los componentes puede encontrarse en el apéndice del capítulo 4 de Sharma (1996).

8 A una correlación fuerte (pero no perfecta) entre dos o más variables independientes se le llama «multicolinealidad» (Wooldridge, 2015, p. 84).

TABLA 4. Coeficientes de correlación de variables étnicas

Variables	(1)	(2)	(3)	(4)
(1) Grupo étnico autoidentificado	1,000			
(2) Lengua materna	0,520***	1,000		
(3) Área rural	0,236***	0,301***	1,000	
(4) Migración	-0,027***	-0,002***	-0,103***	1,000

*** p < 0,01, ** p < 0,05, * p < 0,1

Fuente: Encuesta Nacional de Hogares (2017)

Elaboración propia

Para aplicar un ACP nos encontramos con un primer problema: que esta técnica es aplicada principalmente cuando se poseen variables continuas; sin embargo, en nuestro caso, tenemos variables categóricas, es decir, variables que poseen categorías excluyentes entre sí. Para solucionar esta dificultad, construimos el índice de tal manera que cada variable étnica sea representada en forma dicotómica, es decir, con solo dos categorías que tienen valores de cero o uno. De esta manera, luego de recodificar las variables categóricas que teníamos, podemos proceder con el ACP (Vyas y Kumaranayake, 2006).

Según Kolenikov y Angeles (2009), el uso de variables categóricas con más de dos categorías podría generar correlaciones espurias; por lo tanto, recomiendan que, en vez de usar la correlación de Pearson (que se usa en el ACP estándar), se use la correlación policórica. La correlación policórica asume una variable normal subyacente para cada una de las variables categóricas (Stas Kolenikov, 2016). Por lo tanto, procederemos a hacer un ACP con matriz policórica⁹; los resultados serán igualmente interpretables en caso hubiéramos aplicado un ACP estándar, pero optamos por la primera opción debido a que contamos con variables categóricas y este procedimiento es el más apropiado.

En algunos casos, el ACP es innecesario si las variables iniciales no están lo suficientemente correlacionadas. Para comprobar esto, podemos usar el test de Kaiser-Meyer-Olkin (KMO). El estadístico KMO toma valores entre 0 y 1. Un valor más cercano a 0 implica que las variables iniciales tienen muy poco en común (baja correlación) y no aplicaría un ACP; sin embargo, mientras más cercano sea el valor a 1, sí se puede proceder con un ACP. Como parte de una convención en la literatura aplicada, un valor del KMO mayor a 0,5 nos sugeriría proceder con un ACP. En la tabla 5

9 En la presente investigación se optó por el uso del programa estadístico STATA 16 para generar la matriz de correlación policórica usando el comando «polychoric».

podemos observar que el estadístico KMO en conjunto es 0,59, lo cual nos sugiere que un ACP en este caso sí es adecuado.

TABLA 5. Medida de adecuación muestral Kaiser-Meyer-Olkin

Variable	KMO
(1) Grupo étnico autoidentificado	0,576
(2) Lengua materna	0,561
(3) Área rural	0,691
(4) Migración	0,471
En conjunto	0,586

Luego de confirmar la pertinencia de un ACP, la siguiente pregunta sería cuántos componentes deberíamos considerar. Esta decisión depende de cuánta información estamos dispuestos a sacrificar, lo cual es, por su puesto, una decisión subjetiva (Sharma, 1996); sin embargo, existen reglas sugeridas. Tal como se comentó previamente sobre los componentes principales, de acuerdo con el orden, cada componente suele capturar una proporción de la varianza de las variables iniciales. En esta línea, contar con un gráfico de sedimentación puede ayudarnos en la decisión del número de componentes que deberíamos considerar.

Tal como esperábamos, el primer componente es el que está asociado con una proporción mayor de la varianza de los datos iniciales. Por convención, el número de componentes asociados a valores por encima de 1 (la línea horizontal) son los que deberíamos considerar. En la figura 1 vemos que dos componentes están por encima de la línea horizontal; sin embargo, el segundo componente es muy cercano, por lo que decidimos considerar solamente un componente. Con esto argumentamos que es posible reducir la dimensión de las cuatro variables de etnicidad en un solo indicador.

Como comentamos anteriormente, hemos optado por usar un ACP con matriz de correlación policórica, por su flexibilidad para trabajar con variables dicotómicas. En la tabla 6 observamos algunos estadísticos sobre el índice de etnicidad. Por ejemplo, vemos que, para todos los años, el valor mínimo del índice es de -1,62, mientras que el máximo es 1,28 (excepto en 2017, en que aumenta un decimal). La manera de interpretar este índice es que los valores más bajos se asocian con menor intensidad étnica, a diferencia de los valores más altos. El valor del índice en términos absolutos no tiene una interpretación por sí solo.

Tal como muestra la figura 2, tenemos un diagrama de cajas del índice de etnicidad según el año, en el cual hemos considerado solamente el

primer y último año del periodo que pretendemos analizar. El promedio del índice a nivel nacional en 2007 fue -0,60, mientras que en 2017 fue -0,62, lo que podríamos interpretar como si en un período de 10 años se tuvo una menor intensidad étnica.

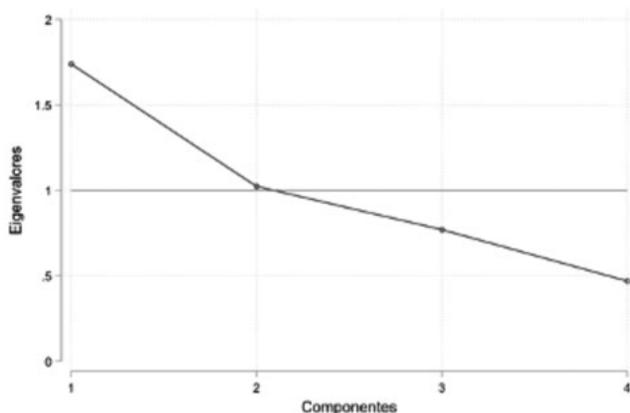


FIGURA 1. Gráfico de sedimentación

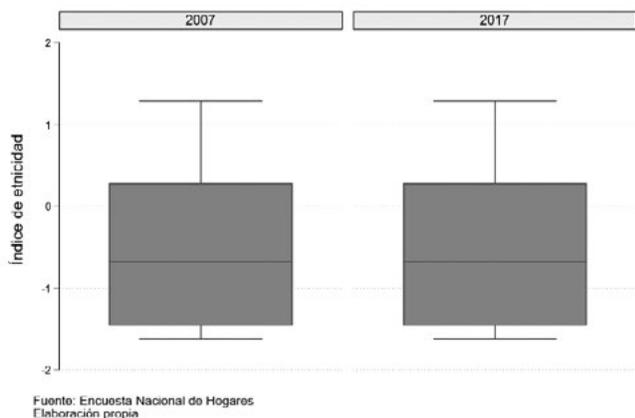


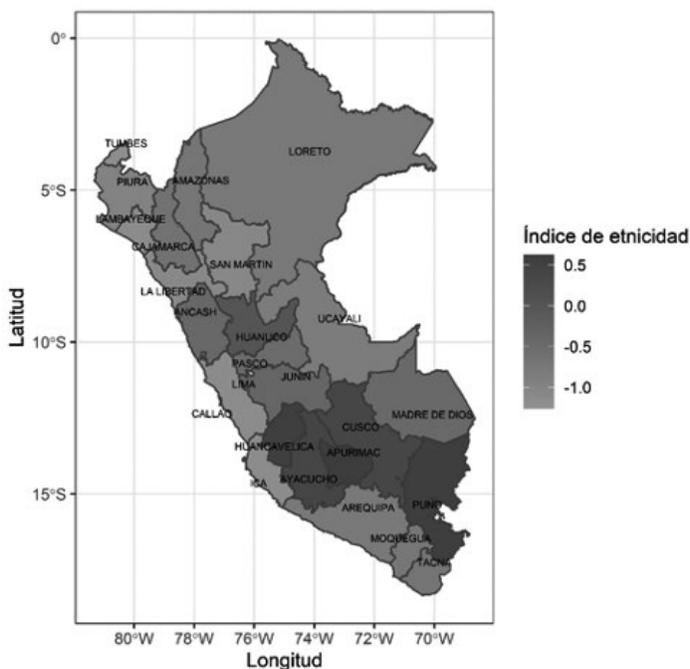
FIGURA 2. Diagrama de cajas del índice de etnicidad

TABLA 6. Estadísticos descriptivos sobre el índice de etnicidad

	2007	2012	2017
Observaciones	77 188	85 992	105 436
Media	-0,60	-0,62	-0,62
Mínimo	-1,62	-1,62	-1,62
Máximo	1,28	1,28	1,29
Desviación estándar	0,95	0,96	0,96

Fuente: Encuesta Nacional de Hogares
Elaboración propia

El índice de etnicidad puede también descomponerse según el departamento, como se muestra para 2017 en la figura 3. En este mapa, los departamentos con una mayor intensidad étnica están con una tonalidad más oscura. Podemos distinguir que, para 2017, departamentos como Huancavelica, Ayacucho, Apurímac, Cusco y Puno tienen mayores índices de etnicidad, a diferencia de Lima, Ica o Lambayeque, cuyos respectivos índices de etnicidad son los de menor valor.



Fuente: Encuesta Nacional de Hogares (2017)
Elaboración propia

FIGURA 3. Mapa departamental del índice de etnicidad (2017)

Dadas las diferencias en los valores promedios del índice de etnicidad entre departamentos y grupos de departamentos con similitud geográfica, en la figura 4 comparamos el índice según dominio geográfico. Lo primero que podemos observar es que, comparando 2007 con 2017, el índice de etnicidad disminuyó en la selva, costa norte, sierra norte y sierra sur, aunque en esta última la reducción fue mínima. Luego, podemos ver que la sierra sur y el centro poseen valores significativamente mayores a los otros dominios. Incluso, tanto para 2007 como para 2017, el índice de etnicidad está por encima del promedio nacional del periodo de análisis. Vemos, entonces, la heterogeneidad del índice según la ubicación geográfica del individuo.

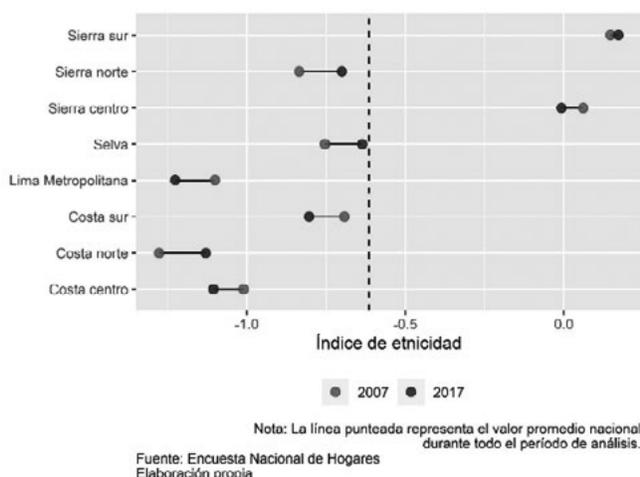


FIGURA 4. Índice de etnicidad según dominio geográfico y año

Resultados

En la tabla 7 se desagrega el índice entre 3 categorías de pobreza. Se encuentra que hay una mayor intensidad étnica si el individuo es pobre extremo en cada año analizado. Asimismo, si es no pobre, existe una menor intensidad étnica; este valor disminuye de 2007 a 2017 en un 14,6 %. A la luz de estos valores, se puede indicar que los grupos étnicos están más expuestos a problemas estructurales como la pobreza, lo cual sería consistente a lo presentado en otros estudios (Baldárrogo, 2017; Pasquier-Doumer y Risso Brandon, 2015; Presidencia del Consejo de Ministros, 2013).

TABLA 7. Índice de etnicidad y pobreza

	2007	2017	Var (%)
Pobre extrema	0,30	0,25	-14,6 %
Pobre no extremo	-0,45	-0,17	-62,6 %
No pobre	-0,92	-0,78	-15,5 %

Fuente: Encuesta Nacional de Hogares

Elaboración propia

La etnicidad también puede asociarse con algunas variables educativas como el analfabetismo, el nivel educativo y los años de educación. A continuación, pasaremos a describir las relaciones del índice de etnicidad con estas variables.

En la figura 5 se observa que el índice es mayor cuando no se cuenta con un nivel educativo para 2007 y 2017. También se puede indicar que la intensidad étnica se reduce a medida que se incrementa el nivel educativo. Por ejemplo, para el nivel universitario el índice va a ser menor, lo cual se correlaciona con una menor presencia de grupos intensamente étnicos en dicho nivel. Estos resultados van en línea con los datos descriptivos de la tabla 2, y se complementan con el modelo de años predichos de escolaridad realizado por Telles *et al.* (2015), en el que se encuentra que la calificación de una piel más oscura se relaciona con menores años de educación para el caso peruano.

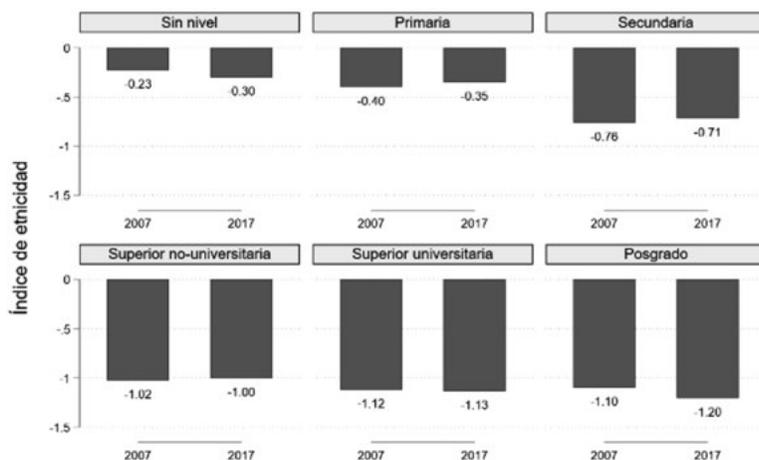
Fuente: Encuesta Nacional de Hogares
Elaboración propia

FIGURA 5. Índice de etnicidad por nivel educativo

En la tabla 8 tenemos que las personas analfabetas tienen un índice de etnicidad mayor, a diferencia de aquellas que sí saben leer y escribir. Si comparamos según año, podemos observar que hubo un aumento en el índice para las personas analfabetas, mientras que para la población que sí sabe leer y escribir hubo una reducción en el índice.

TABLA 8. Índice de etnicidad y analfabetismo

	2007	2017	Var (%)
Sabe leer y escribir	-0,73	-0,71	-1,7 %
No sabe leer ni escribir	0,19	0,30	58,9 %

Nota: Se consideraron a mayores de 15 años

Fuente: Encuesta Nacional de Hogares

Elaboración propia

Además, resulta interesante analizar la relación entre el índice de etnicidad y los años de educación alcanzados. En la figura 6 representamos esta relación con un gráfico de dispersión¹⁰ y encontramos que existe una correlación negativa entre estas variables. Esto significa que mayores valores del índice de etnicidad se asocian con menores años de educación alcanzados. Esta relación sigue siendo la misma para 2007 y 2017. Sin embargo, la situación parece haber mejorado entre ambos años, ya que, dado un valor del índice, tendríamos mayores años de educación en 2017 que en 2007. Estos resultados van en línea con la relación entre el logro educativo, el desarrollo cognitivo y los grupos étnicos minoritarios mostrada por Arouri *et al.* (2019).

El autoempleo es una situación laboral en la que una persona tiene la necesidad de generar ingresos mediante la creación de un puesto de trabajo propio. En el Perú existe un gran porcentaje de individuos auto-empleados en áreas urbanas, que deciden voluntariamente pertenecer al sector informal, pues obtienen ganancias mayores a las del sector formal (Saavedra y Chong, 1999; Yamada, 1996). Sobre la etnicidad y el autoempleo, en la tabla 9 observamos que, para todos los dominios geográficos, la población autoempleada tiene un índice de etnicidad mayor, lo que acontece también a nivel nacional. Las personas con una mayor intensidad étnica están más propensas a estar en una situación de autoempleo, lo que se relacionaría con la discriminación étnica existente en el mercado laboral. Esto ya ha sido estudiado por otros autores en el Perú, como

¹⁰Específicamente, usamos el comando «binscatter» en Stata, el cual nos brinda una forma no paramétrica de visualizar la relación entre dos variables cuando tenemos una larga cantidad de datos (Stepner, 2013).

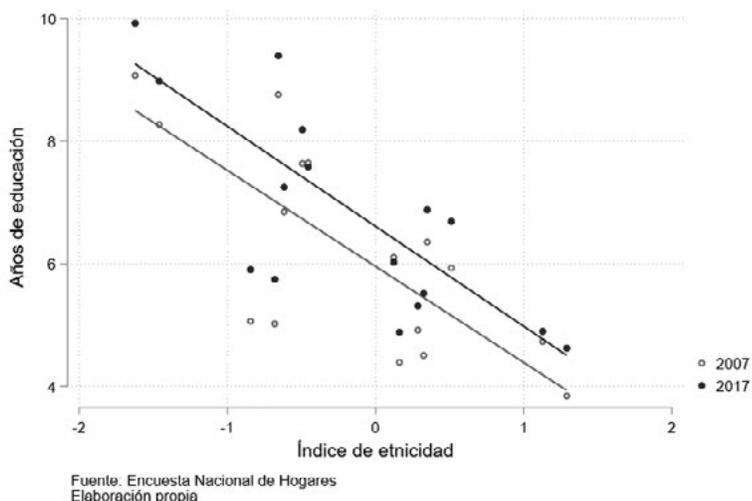


FIGURA 6. Índice de etnicidad y años de educación

Galarza y Yamada (2012), quienes mostraron la discriminación por parte de los empleadores que existiría en el mercado laboral formal, lo cual podría empujar, por necesidad, a que las personas con una mayor intensidad étnica busquen alguna forma de generar ingresos.

TABLA 9. Índice de etnicidad según autoempleo y dominio geográfico (2007-2017)

Dominio geográfico	No autoempleo	Autoempleado (empleado u obrero)	Total
Costa norte	-1,28	-1,22	-1,25
Costa centro	-1,11	-1,07	-1,09
Costa sur	-0,79	-0,52	-0,69
Sierra norte	-1,05	-0,77	-0,88
Sierra centro	-0,32	0,25	0,00
Sierra sur	-0,20	0,51	0,18
Selva	-0,96	-0,64	-0,79
Lima Metropolitana	-1,25	-1,15	-1,22
Nacional	-0,87	-0,42	-0,66

Fuente: Encuesta Nacional de Hogares
Elaboración propia

Es de esperarse que, al tener la etnicidad este vínculo con la condición laboral, exista una relación también con el ingreso laboral. En la figura 7 se aprecia una correlación negativa entre el índice de etnicidad y el ingreso

laboral mensual para 2007 y 2017. Dado un valor del índice de etnicidad, el ingreso laboral mensual será mayor en 2017 que en 2007. A pesar de esto, a valores más intensos de etnicidad, la brecha se va cerrando, lo que significa que, para individuos con una mayor intensidad étnica, sus ingresos actuales no se diferenciarían sustancialmente a los de hace 10 años. Al ser el ingreso laboral un elemento clave para el consumo y, por tanto, el bienestar de un hogar, los grupos que tienen más intensidad étnica son propensos a quedarse estancados en una trampa de la pobreza.

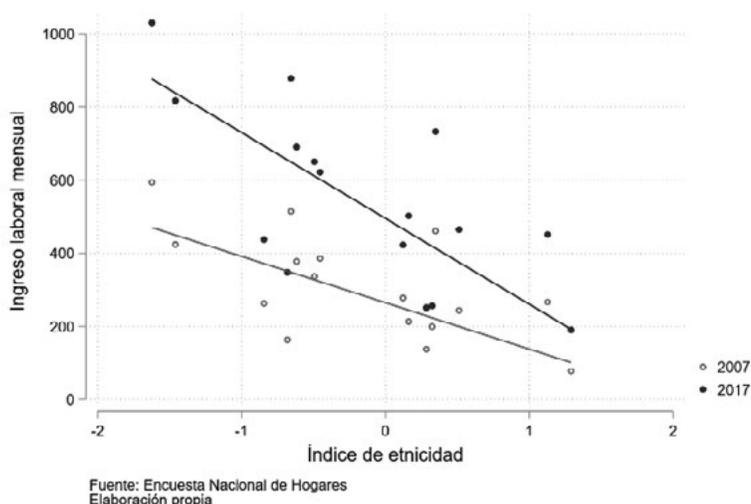


FIGURA 7. Índice de etnicidad e ingreso laboral mensual (2007-2017)

Discusión

En el Perú conviven numerosos grupos étnicos que manifiestan sus respectivas culturas de distintas formas. En la presente investigación hemos identificado características personales asociadas a la etnicidad y construido un índice que capture todas estas y las resuma en un solo indicador. Encontramos que mayores valores del índice, que podemos asociar con mayor intensidad de estas características étnicas, se encuentran en determinados geográficos, como la sierra sur y central, lo cual se mantiene a lo largo de todo el periodo de análisis. Esto nos motivó a que, a través del índice propuesto, exploremos las relaciones existentes entre la etnicidad y variables educativas y laborales.

Por el lado educativo, los resultados nos muestran que, en los niveles educativos más altos, el valor del índice es menor, es decir, habría menores personas con características étnicamente diversas. Además, existe una mayor etnicidad en la población que es analfabeta; la diferencia del índice es considerable respecto a la población que sí sabe leer y escribir. Por el lado laboral, un mayor valor del índice de etnicidad se correlaciona con menores ingresos y con la posibilidad de encontrarse autoempleado.

Si bien la investigación nos muestra resultados interesantes sobre la etnicidad y las oportunidades educativas y laborales, hemos presentado solamente un análisis correlacional. Debemos reconocer que la correlación entre dos variables no implica causalidad; por tanto, los resultados deben ser interpretados bajo esa premisa. Asimismo, para investigaciones futuras, se recomienda seguir explorando las relaciones de la etnicidad con, por ejemplo, el acceso a la salud, el grado de ciudadanía y la gobernabilidad, usando modelos econométricos en los que se vinculen estas variables. Debido a que el Perú es un país con una gran diversidad étnica, queda pendiente en la agenda de investigación seguir explorando las brechas existentes en educación, ingresos o salud, para tener un mejor panorama que nos ayude mejorar la inclusión de estos grupos étnicos.

Contribución de autoría

Ronny Condor y Guiliana Campos contribuyeron igualmente en la elaboración del marco teórico, el análisis de los resultados y la redacción de la versión final del presente artículo.

Fuente de financiamiento

Ninguno.

Potenciales conflictos de interés

Ninguno.

REFERENCIAS BIBLIOGRÁFICAS

- Ahmad, W. I. U. y Bradby, H. (2007). Locating ethnicity and health: exploring concepts and contexts. *Sociology of Health & Illness*, 29(6), pp. 795-810.
- Alesina, A., Michalopoulos, S. y Papaioannou, E. (2016). Ethnic inequality. *Journal of Political Economy*, 124(2), pp. 428-488. <https://doi.org/10.1086/685300>
- Arouri, M., Ben-Youssef, A. y Nguyen, C. V. (2019). Ethnic and racial disparities in children's education: Comparative evidence from Ethiopia, India, Peru and Viet Nam. *Children and Youth Services Review*, 100, pp. 503-514.
- Baldárrago, E. (2017). Dinámicas étnicas en el Perú: hacia una caracterización y tipología para el diseño de las políticas públicas. Instituto Nacional de Estadística e Informática, INEI.
- Bello M., A. y Rangel, M. (2000). *Etnicidad, «raza» y equidad en América Latina y el Caribe*. Comisión Económica para América Latina y el Caribe, Cepal.
- Burton, J., Nandi, A. y Platt, L. (2010). Measuring ethnicity: challenges and opportunities for survey research. *Ethnic and Racial Studies*, 33(8), pp. 1332-1349. <https://doi.org/10.1080/01419870903527801>
- Figuroa, A. y Barrón, M. (2005). *Inequality, ethnicity and social disorder in Peru*. Centre for Research on Inequality, Human Security and Ethnicity, CRISE.
- Galarza, F. y Yamada Fukusaki, G. (2012). Discriminación laboral en Lima: el rol de la belleza, la raza y el sexo. <https://repositorio.up.edu.pe/handle/11354/388#.YWZYIQ7fiSY.mendeley>
- Geertz, C. (1993). *The Interpretation of cultures*. Fontana.
- Giraldo, M. A., Díaz, A. Q. y Calderón, M. R. (2019). Migración, autoidentificación y religión: los estudiantes de la Facultad de Letras y Ciencias Humanas (Universidad Nacional Mayor de San Marcos). *Praxis*, 15(2), pp. 207-221.
- Huaman, M. G., Silva, G. M. y Sulmont, D. (2021). Coproduciendo categorías étnico-raciales: empadronadores en el Censo Nacional de Perú 2017. *Debates en Sociología*, 53, pp. 1-37.
- Instituto Nacional de Estadística e Informática, INEI. (2017). *La autoidentificación étnica: población indígena y afroperuana*. INEI. https://www.inei.gob.pe/media/MenuRecursivo/publicaciones_digitales/Est/Lib1642/
- Kolenikov, S. y Angeles, G. (2009). Socioeconomic status measurement with discrete proxy variables: is the Principal Component Analysis a reliable answer? *Review of Income and Wealth*, 55(1), pp. 128-165. <https://doi.org/10.1111/j.1475-4991.2008.00309.x>
- Kolenikov, S. (2016). Polychoric, by any other «namelist». Stata Conference.

Kolo, P. (2012). *Measuring a new aspect of ethnicity: The appropriate diversity index*. <https://www.econstor.eu/handle/10419/92997>

Martel, P., Mbofana, F. y Cousens, S. (2021). The polychoric dual-component wealth index as an alternative to the DHS index: Addressing the urban bias. *Journal of Global Health, 11*, pp. 1-19. <https://doi.org/10.7189/JOGH.11.04003>

Martin, E. A. y Gerber, E. (2006). Methodological influences on comparability of race measurements: Several cautionary examples. *Survey Methodology, 6*.

McDonnell, J. y De Lourenco, C. (2009). You're Brazilian, right? What kind of Brazilian are you? The racialization of Brazilian immigrant women. *Ethnic and Racial Studies, 32*(2), pp. 239-256.

Morley, S. (2017). Changes in rural poverty in Perú 2004-2012. *Latin American Economic Review, 26*(1), pp. 1-20.

Ñopo, H., Saavedra, J. y Torero, M. (2007). Ethnicity and earnings in a mixed-race labor market. *Economic Development and Cultural Change, 55*(4), pp. 709-734. <https://doi.org/10.1086/516762>

Paredes, M. y House, Q. E. (2007). *Fluid identities: Exploring ethnicity in Peru*. Centre for Research on Inequality, Human Security and Ethnicity, CRISE.

Pasquier-Doumer, L. y Risso Brandon, F. (2015). Aspiration failure: A poverty trap for indigenous children in Peru? *World Development, 72*, pp. 208-223. <https://doi.org/10.1016/J.WORLDDEV.2015.03.001>

Phinney, J. S. (1990). Ethnic identity in adolescents and adults: review of research. *Psychological Bulletin, 108*(3), p. 499.

Presidencia del Consejo de Ministros. (2013). Perú: Tercer Informe Nacional de Cumplimiento de los Objetivos de Desarrollo del Milenio. Organización de las Naciones Unidas. Sistema de Naciones Unidas en el Perú.

Ravallion, M. (2019). *Ethnic inequality and poverty in Malaysia since 1969*. <https://doi.org/10.3386/W25640>

Roth, W. D. (2016). The multiple dimensions of race. *Ethnic and Racial Studies, 39*(8), pp. 1310-1338. <https://doi.org/https://doi.org/10.1080/01419870.2016.1140793>

Saavedra, J. y Chong, A. (1999). Structural reform, institutions and earnings: Evidence from the formal and informal sectors in urban Peru. *The Journal of Development Studies, 35*, pp. 95-116. <https://doi.org/https://doi.org/10.1080/00220389908422582>

Saavedra, J., Torero, M. y Ñopo, H. (2004). *Ethnicity and earnings in urban Peru*. Institute for the Study of Labor Discussion Papers, 980.

- Schkolnik, S. (2009). La inclusión del enfoque étnico en los censos de población de América Latina. *Notas de Población, Cepal*, 89, pp. 57-100. <http://hdl.handle.net/11362/12857>
- Sharma, S. (1996). *Applied multivariate techniques*. John Wiley & Sons Inc.
- Stepner, M. (2013). BINSCATTER: Stata module to generate binned scatter-plots. EconPapers.
- Sulmont, D. (2010). Raza y etnicidad desde las encuestas sociales y de opinión: dime cuántos quieres encontrar y te diré qué preguntar... <https://centroderecursos.cultura.pe/sites/default/files/rb/pdf/Raza%20y%20etnicidad%20desde%20las%20encuestas%20sociales%20y%20de%20opin-ion.pdf>
- Sulmont, D. (2011). Race, ethnicity and politics in three Peruvian localities: An analysis of the 2005 CRISE perceptions survey in Peru. *Latin American and Caribbean Ethnic Studies*, 6(1), pp. 47-78. <https://doi.org/10.1080/17442222.2011.543873>
- Taş, E. O., Reimão, M. E. y Orlando, M. B. (2014). Gender, ethnicity, and cumulative disadvantage in education outcomes. *World Development*, 64, pp. 538-553. <https://doi.org/10.1016/J.WORLDDEV.2014.06.036>
- Telles, E., Flores, R. D. y Urrea-Giraldo, F. (2015). Pigmentocracies: Educational inequality, skin color and census ethnorracial identification in eight Latin American countries. *Research in Social Stratification and Mobility*, 49, pp. 39-58. <https://doi.org/10.1016/J.RSSM.2015.02.002>
- Trivelli, C., Von Hesse, M., Diez, A. y Del Castillo, L. (2000). *Desafíos del desarrollo rural en el Perú*. Consorcio de Investigación Económica y Social.
- Vyas, S. y Kumaranayake, L. (2006). Constructing socio-economic status indices: how to use principal components analysis. *Health Policy and Planning*, 21(6), pp. 459-468. <https://doi.org/10.1093/heapol/czl029>
- Woo-Mora, L. G. (2021). Unveiling the cosmic race: Racial inequalities in Latin America. *SSRN Electronic Journal*. <https://doi.org/10.2139/SSRN.3870741>
- Wooldridge, J. M. (2015). *Introductory econometrics: A modern approach*. Cengage Learning.
- Yamada, G. (1996). Urban informal employment and self-employment in developing countries: Theory and evidence. *Economic Development and Cultural Change*, 44(2), pp. 289-314. <https://doi.org/https://doi.org/10.1086/452214>

Recepción: 15/10/2021
Aceptación: 27/2/2022