

Alternativas de estandarización para índices compuestos espacio-temporales. El caso del rezago educativo en los estados de México, 2000 a 2020

Standardization alternatives for space-time composite indices. The case of educational backwardness in the states of Mexico from 2000 to 2020

Jesús Antonio Treviño Cantú*

Recibido: 26/07/2022. Aprobado: 23/09/2022. Publicado: 25/11/2022.

Resumen. Los investigadores de índices compuestos han detectado y tratado de eliminar, con poco éxito hasta ahora, la ponderación implícita de los datos porque generan índices distorsionados. El origen de esta distorsión suele adjudicarse a las variables con distinto rango de valores, al valor máximo y mínimo desigual entre ellas o a la de asimetría de cada variable. El procedimiento para abordar este triple reto es la estandarización de los datos. A la fecha no hay un procedimiento de este tipo que corrija simultáneamente este triple problema y, con ello, la ponderación implícita. El igual máximo y mínimo entre variables resuelve el problema de rangos desiguales, pero no corrige la asimetría en cada una de ellas. El rango igual entre variables no implica igual máximo y mínimo entre ellas ni garantiza simetría en cada variable. Finalmente, la corrección de la asimetría en cada variable no implica igual máximo y mínimo ni igual rango entre variables.

El objetivo del presente estudio es sugerir e ilustrar un procedimiento de estandarización de variables que elimina la ponderación implícita de cada variable. Las estandarizaciones revisadas en este trabajo para elaborar índices compuestos resuelven solo de forma parcial el problema de la ponderación implícita por lo que este trabajo propone un procedimiento de estandarización que resuelve simultáneamente el triple reto impuesto por los rangos dispares, los máximos y mínimos desiguales y la asimetría. El nuevo procedimiento propuesto, la estandarización balanceada, soluciona el pro-

blema de la ponderación implícita en el tiempo y el espacio, en la agregación transversal y longitudinal de las variables de las distintas unidades de observación. Este trabajo ilustra el procedimiento de estandarización propuesto con las variables del rezago educativo en los estados de México. La metodología propuesta es aplicable a cualquier tema donde la estandarización espacio temporal sea necesaria.

Palabras clave: procedimientos de estandarización, ponderación implícita, rangos desiguales, valores máximo y mínimo distintos, asimetría.

Abstract. Researchers of composite indexes have detected and attempted to eliminate the implicit weighting of data because it leads to biased indexes, with little success so far. This bias is usually attributed to variables having different ranges of values, unequal maximum and minimum values between variables, or the asymmetry of each variable. The procedure for addressing this triple challenge is data standardization. No procedure is currently available that simultaneously solves this triple issue and, with it, implicit weighting. Equal maximum and minimum values between variables solve the unequal-range issue but do not correct the asymmetry of each variable. An equal range between variables does not imply equal maximum and minimum

* Universidad Autónoma de Nuevo León. Facultad de arquitectura, Postgrado en asuntos urbanos. Emilio Carranza 2618, col. Lomas de Tolteca, 67170, Ciudad Guadalupe, Nuevo León. ORCID: <https://orcid.org/0000-0002-0211-2140>. Email: jtrevino41@hotmail.com

values between them, nor does it guarantee the symmetry of each variable. Finally, correcting the asymmetry of each variable does not imply equal maximum and minimum values or equal ranges between variables.

The objective of this study is to suggest and illustrate a variable standardization procedure that eliminates the implicit weighting of each variable. The standardizations reviewed in this paper to elaborate composite indexes solve only partially the implicit weighting issue. Thus, this paper proposes a standardization procedure that simultaneously solves the triple challenge imposed by unequal ranges, unequal maxima and minima, and asymmetry. The proposed

new procedure, balanced standardization, solves the issue of implicit weighting in time and space in the cross-sectional and longitudinal aggregation of variables of the different observation units. This work illustrates the proposed standardization procedure with variables related to the educational lag in the states of Mexico. The methodology proposed herein is applicable to any topic requiring spatio-temporal standardization.

Keywords: standardization procedures, implicit weighting, unequal ranges, different minimum and maximum values, asymmetry.

INTRODUCCIÓN

Un índice compuesto (IC) es la manera sintética de expresar numéricamente una idea o mensaje complejo sobre un problema específico sin perder de vista el marco general del fenómeno considerado. Este valor sintético solo tiene sentido cuando se compara con otros valores en un lugar, en un momento dado (comparación transversal), o consigo mismo en el tiempo (comparación longitudinal). La comparación transversal o longitudinal es posible gracias a la estandarización de las variables.

Los datos originales, por ejemplo, pueden estar disponibles en porcentajes, kilómetros, personas por cuarto o kilómetro cuadrado, minutos, microgramos por metro cúbico, tasas por mil o cien mil. La estandarización expresa las variables originales en una misma unidad de medida abstracta o general a fin de que sean comparables y agregadas en un IC. Otras funciones de la estandarización son invertir la polaridad negativa de las variables de un IC, controlar los pesos implícitos en el rango de valores y la forma de la distribución, identificar los casos atípicos que dificultan el uso de la estadística paramétrica, y asegurar que las variables sean comparables en el tiempo y espacio. Los procedimientos de estandarización abordados en este escrito cubren estos aspectos, con especial énfasis en tres características que ejercen un peso implícito al agregar variables en un IC: el rango desigual de los valores, los valores máximo y mínimo desiguales entre variables, y la asimetría de cada distribución.

Rango desigual. El rango es la distancia entre el valor máximo y mínimo de la variable. El rango desigual entre variables produce índices distorsio-

nados porque tienen sesgo hacia los valores mayores (ponderación implícita).¹ En las distribuciones de rango desigual domina la mayor amplitud porque los valores bajos en unas variables son relativamente altos en otras. Ejemplo: en la 'estandarización mín-máx restringida' (Z_{MP} en esta investigación) de Mazziotta y Pareto (2022) las variables tienen igual rango, pero el empate de valores en los extremos y la asimetría quedan como problemas sin resolver.

Empate de valores máximo y mínimo. Las variables con mayores valores en los extremos dominan en el índice agregado, aunque tengan el mismo rango. Los valores bajos en variables con valores altos en los extremos (por ejemplo, 50 a 80) parecerán altos para variables con valores bajos en los extremos (5 a 35), aunque haya igual amplitud en las variables (30 en el ejemplo). Ejemplo: en la estandarización mín-máx (Z_{MM} en este estudio), explicado en la OECD (2008), empata los valores en los extremos en variables de igual rango sin modificar la asimetría de los valores originales.

Asimetría. Las variables con igual rango e iguales valores máximo y mínimo pueden tener distintos grados de asimetría. La asimetría crea o magnifica el efecto sustitución o compensatorio en la agregación lineal de las variables. Hay dos tipos de asimetría, a la derecha o a la izquierda. La primera también es conocida como asimetría positiva, muestra la concentración de casos en los valores bajos creando una distribución de cola a la derecha. La media aparece a la derecha de (superior a) la mediana de manera

¹ El 'rango' de una distribución de valores continuos equivale al range (en inglés), distinto de rank, que implica jerarquía de carácter ordinal.

tal que la suma o promedio de las variables con mayor valor compensan o ‘esconden’ las variables con menor valor. En las variables con asimetría a la izquierda (asimetría negativa) la mayoría de los casos tiene valores altos en la variable original. En este segundo caso, las variables no diferencian las observaciones significativamente y deben ser omitidas en la investigación (Balcerzak, 2016).

La corrección de la asimetría usualmente ajusta también a la curtosis porque la normalización suaviza y obtiene una distribución simétrica, sin picos exagerados o datos aglomerados en los extremos. Esta corrección puede verificarse con la curtosis estandarizada (zcurtosis), cuyo valor, como en el sesgo estandarizado (zsesgo),² debe estar en el rango ± 2 .

Hay dos opciones para corregir la asimetría (y la curtosis) de las variables. La primera opción, la tradicional, incluye dos pasos (Gilthorpe, 1995): i) ‘normalizar’ las variables mediante los procedimientos de transformación conocidos, ya sea logs, raíz cuadrada, inversa de la variable, Box-Cox o el arco seno de la variable, y ii) la variable normalizada es estandarizada por un procedimiento usual, como el método mín-máx aplicado al Índice de Desarrollo Humano (IDH). La estandarización mín-máx logra igual rango y empate de valores en los extremos de los datos previamente normalizados.³

La segunda opción para corregir la asimetría consiste en transformar y estandarizar de forma simultánea la variable, como en z_{MAD} o en tres procedimientos en esta investigación. No es necesario que los procedimientos normalicen los datos, solo que sus rangos sean iguales, los valores en los extremos también sean iguales, y la asimetría estandarizada sea aceptable (± 2).

La regla no escrita en los IC es que las variables sean positivas, que un incremento en la variable se refleje en un aumento en el IC. Si no fuera así, cada procedimiento de estandarización debe

proporcionar la opción para invertir la polaridad de la variable.

Planteamiento del problema. Hasta ahora no hay un procedimiento que neutralice simultáneamente el peso implícito en la asimetría en cada variable, los rangos desiguales entre variables y los valores máximo y mínimo entre ellas.

Objetivo general. Proponer e ilustrar un procedimiento que controle la asimetría, genere valores iguales en los extremos superior e inferior y, por tanto, igual rango entre variables. La revisión y aplicación de los procedimientos actuales proporciona el marco metodológico y empírico al método propuesto. La investigación ilustra la aportación metodológica del nuevo procedimiento con las variables del rezago educativo que son parte del índice de rezago social en las regiones y los estados de México en el periodo 2000 a 2020.

Los IC son usuales en diversos campos disciplinarios. El procedimiento sugerido de estandarización balanceada es aplicable a variables continuas de índices compuestos de todo tipo, elaborados para distinta escala geográfica y periodo de tiempo.

METODOLOGÍA

La investigación examina versiones transversales y longitudinales de procedimientos de estandarización alternativos. Las versiones transversales de estos índices son primero evaluadas con diagramas de caja y estadística descriptiva básica para seleccionar posteriormente las versiones longitudinales aplicables al estudio de caso.

Los procedimientos de estandarización pueden ser transversales (relativos), si refieren una distribución espacial en un momento en el tiempo (por ejemplo, un año), o longitudinales (absolutos), si consideran una sucesión de momentos (por ejemplo, varios años de un periodo).

La investigación se concentra en índices transversales que tienen también versiones longitudina-

² El zsesgo=Valor de la asimetría/Error de la asimetría. La fórmula de la zcurtosis es similar, remplazando la asimetría por la curtosis.

³ En esta investigación la ‘normalización’ o ‘transformación’ es una operación distinta a la ‘estandarización’. La transformación modifica la forma de la distribución por lo que es apropiada para eliminar la asimetría. La estandarización convierte los datos en unidades comparables, si fuera necesario,

con igual rango e igual valores en los extremos sin modificar la forma de la distribución (excepto en la estandarización balanceada sugerida en esta investigación)

les. Esta decisión descarta los índices de naturaleza ordinal o jerárquica que alteran la distancia entre los casos, tal como Knox (Ricketts, Ilbery y Kneafsey, 2006), Borda (Dasgupta y Weale, 1992), o percentil (Acharya y Porwal, 2020; Flanagan, Gregory, Hallisey, Heitgerd, y Lewis, 2011). Estos índices son de uso exclusivo transversal. La única comparación posible espacio temporal de los índices relativos refiere los desplazamientos entre los rangos, sin que ello implique incrementos o disminuciones absolutas en el tiempo de la variable estandarizada. La investigación también omite la indización o ‘distancia a una referencia’ o valor de un año específico por no modificar el coeficiente de variación de los valores originales (Mazziotta y Pareto, 2022 y 2021).

Estandarización transversal

a) *Estandarización clásica* (z_{t0}). Los datos originales usualmente tienen medidas de tendencia central distintas (p. ej., medias y medianas), amplitud de valores desiguales y distribuciones asimétricas. La estandarización usualmente interesa como artificio para expresar las variables en una misma unidad de medida, sin atender el resto de las características importantes en la elaboración de índices compuestos. Este procedimiento centra la media en cero y la desviación estándar en la unidad. La estandarización z_{t0} solo es apropiada para datos originales normales donde la asimetría no distorsiona el significado de los parámetros (Smith, 1975). La estandarización clásica no elimina la asimetría, los valores mínimo y máximo entre variables son distintos y los rangos entre ellas son desiguales.

$z_{t0} = \frac{x_i - \bar{x}}{\sigma}$, donde x_i = valor que se estandariza; \bar{x} y σ son la media y desviación estándar de la variable.

b) *Cajas estandarizadas* (z_Q). Sibley (1987) sugiere una estandarización en dos pasos para resolver el problema de medianas desiguales. El primero consiste en centrar todos los diagramas de caja en cero para visualizar la dispersión asimétrica de valores entre las variables. Este paso resta la mediana al valor original (‘Cajas centradas’, z_C). La intención de empatar las medianas entre variables es proporcionar una misma referencia a los procedimientos que buscan controlar la dispersión de los valores.

El segundo paso es similar a la obtención de valores z mediante la fórmula clásica $z = (x_i - \text{media}) / \text{desviación estándar}$, pero reemplaza el numerador por la ‘Caja centrada’ y divide entre el rango intercuartílico. El resultado son las ‘cajas estandarizadas’ (z_Q). Este proceso, de forma similar a z_{t0} centra las medianas en cero y re-escala los valores sin empatar los valores en los extremos, eliminar los rangos desiguales entre las variables, ni modificar la asimetría original en cada una de ellas.

Cajas centradas: $z_c = x_i - Me$, donde Me = mediana.

Cajas estandarizadas (z_Q): $z_Q = \frac{x_i - Me}{Q_3 - Q_1}$, donde Q_1 y Q_3 = primer y tercer cuartil, respectivamente.

c) *Estandarización mín-máx* (z_{MM}). Este procedimiento lleva inicialmente los datos a la escala de 0 a 1. Este método empata los valores máximo y mínimo e iguala la amplitud entre variables, pero la asimetría en la distribución persiste. La fórmula expresa los valores en el rango 70 a 130 para una mejor comparación con otras versiones en esta investigación.

$z_{MM} = \left(\frac{x_i - \text{Mín}}{\text{Máx} - \text{Mín}} * 60 \right) + 70$, donde Mín y Máx son el valor mínimo y máximo de la variable, respectivamente.

d) *Estandarización mín-máx restringida* (z_{MP}). Los autores solo presentan la versión longitudinal de este método, pero puede ser fácilmente adaptado al análisis transversal (Mazziotta y Pareto, 2022; Cutillo, Mazziotta y Pareto, 2021). El procedimiento centra su interés en el rango igual entre variables y la menor varianza de la distribución (comparada con z_{MM}), sin empatar el máximo y el mínimo entre ellas ni eliminar la asimetría de cada distribución. Como en z_{MM} , la estandarización z_{MP} elimina el rango desigual entre variables sin garantizar simetría en cada una de ellas. El procedimiento resuelve el problema de los rangos desiguales y, con ello, parcialmente el problema de la varianza desigual. Desafortunadamente, la varianza también está relacionada con la asimetría puesto que la distribución de valores en algunas variables puede estar más cargada hacia un extremo que al otro dentro de un mismo rango. El método es similar a z_{MM} , pero la referencia no es el valor mínimo, sino un dato conveniente a la interpretación (por ejemplo, promedio, mediana o valor en un caso emblemático).

$z_{MP} = \left(\frac{x_i - Ref}{Máx - Mín} * 60 \right) + 70$, donde $Ref =$ mediana o media.

e) *Normalización robusta* (z_{NR}). Este método empata las medianas y el tamaño de las cajas en un recorrido intercuartílico de 1 a -1 (Brimicombe, 2009). El procedimiento z_{NR} centra la atención en $Q1$, Me , y $Q3$ para estandarizar y normalizar los valores, pero los extremos máximo y mínimo quedan indefinidos. La asimetría es controlada sin que el problema de amplitud desigual entre las variables sea resuelto. $Q1$ y $Q3$ son llevados a la escala de 70 a 130 para fines comparativos. La mediana es escalada a 100.

$$z_{NR} = \left(\frac{x_i - Me}{Me - Q1} * 30 \right) + 100, \text{ para } x_i < Me$$

$$z_{NR} = \left(\frac{x_i - Me}{Q3 - Me} * 30 \right) + 100, \text{ para } x_i > Me$$

Términos definidos en fórmulas previas.

f) *Estandarización balanceada* (z_{EB}). Mediana igual a 100 con el máximo y mínimo igual a 130 y 70, respectivamente. En la misma línea de Brimicombe (2000) y Sibley (1987), esta investigación lleva los valores en los extremos de -1 (mínimo) a 1 (máximo), con la mediana centrada en cero. Los problemas de asimetría y rangos desiguales son resueltos; los valores en los extremos superior e inferior son iguales para todas las variables. Los valores de z_{EB} son llevados a la escala 100 ± 30 para comparaciones en la misma escala de z_{MB} , z_{MM} y z_{NR} .

$$z_{EB} = \left(\frac{x_i - Me}{Me - Mín} * 30 \right) + 100, \text{ para } x_i < Me$$

$$z_{EB} = \left(\frac{x_i - Me}{Máx - Me} * 30 \right) + 100, \text{ para } x_i > Me$$

Términos definidos en fórmulas previas.

En el caso de variables con polaridad negativa (no consideradas en el estudio de caso) hay varias opciones: i) multiplicar por -1 el término entre paréntesis; ii) restar la variable estandarizada de la siguiente manera: $200 - z_{EB}$, z_{MP} , z_{MM} , o z_{NR} , y iii) manipular los términos en la fórmula. Una vez realizada la estandarización para polaridad positiva, la opción ii) es la forma más sencilla y rápida para obtener la estandarización de variables con polaridad negativa.

Estandarización longitudinal

La manera en que se estandarizan los datos determina si el índice compuesto es comparable en el tiempo (Rinner y Pietropaolo, 2021). Los datos originales se refieren a las observaciones de cada año por lo que las comparaciones en el tiempo no son factibles, a menos que la comparación sea ordinal (cambios de jerarquía o estrato), descartada en esta investigación. Una opción para superar esta limitación es apilar los cortes transversales de un año bajo otro en una sola base de datos, como lo sugieren Norman (2010 y 2015) y Exeter, Boyle, y Norman (2011). La identificación de los valores máximo y mínimo, mediana y rango en esta base de datos apilados permiten la estandarización de cada variable considerando todos los casos en todos los años a la vez. A este procedimiento se denomina 'estandarización del conjunto' (*joint standardization*, en los términos de Heinrich, Huber, y Rinner, 2016).

La extensión longitudinal de cada índice transversal es la siguiente (nótese que z_Q evoluciona hacia la estandarización robusta y la estandarización balanceada):

Estandarización z_{t0} . Mazziotta y Pareto (2022) y la OECD (2008) sugieren tomar como referencia la media y desviación estándar del año 0, aunque la distribución transversal inicial o longitudinal no sea normal. Estos parámetros se mantienen constantes para todo el periodo. Una versión alternativa es tomar como referencia la media y desviación de los datos apilados para todo el periodo (Norman, 2010 y 2015; Exeter, Boyle y Norman, 2011). Esta investigación no considera esta segunda versión para evitar redundancia y abreviar procedimientos.

Estandarizaciones z_{MM} , z_{MB} , z_{NR} y z_{EB} . Los valores máximo y mínimo corresponden a los datos apilados. En todas las estandarizaciones, excepto z_{MM} , el valor de referencia puede ser un año específico, inicial (por ejemplo, 2000) o intermedio (puede ser 2010 en el periodo 2000 a 2020), que permanezca fijo para los cálculos del periodo. El ejemplo de esta investigación utiliza como referencia la mediana de todos los años apilados. El uso de la mediana en vez de la media es apropiado porque es robusta a la asimetría y su diferencia con la media no es significativa en distribuciones normales. La media,

en cambio, demanda el requisito de normalidad. La estandarización z_Q se calcula de manera similar a estos procedimientos, pero el primer y tercer cuartil toman el lugar del valor máximo y mínimo.

El valor estandarizado de la estandarización balanceada (z_{EB}) está en la escala 100 ± 30 para fines comparativos. Una variante de la z_{EB} sería tomar como referencia (*Ref*) el promedio de un año específico, como lo sugieren Mazziotta y Pareto (2022) para z_{MP} . La *Ref* también podría ser el valor en cada uno de los casos (por ejemplo, estados) del año de referencia (por ejemplo, 2010), en vez del promedio, pero eso implicaría la pérdida de la jerarquía de las observaciones en dicho año (todos los casos del 2010 serían igual a 100). En esta investigación, la *Ref* para z_{MP} y z_{EB} es la mediana del periodo 2000 a 2020.

RESULTADOS

El objetivo de esta sección es ilustrar los procedimientos de estandarización transversal y longitudinal con variables del rezago educativo en las regiones y los estados de México. El rezago educativo es definido por tres variables (componentes, en la terminología del CONEVAL) de uno de los cinco indicadores (dimensiones) del rezago social en México⁴: Población de 15 años o más analfabeta (x1), Población de 6 a 14 años que no asiste a la escuela (x2), y Población de 15 años y más con educación básica incompleta (x3). Los datos provienen del CONEVAL (2021). Los estados son agrupados en las tres grandes regiones de México identificadas por Angoa, Pérez-Mendoza, y Polèse (2009).

El estudio transversal incluye dos tareas. La primera presenta las características principales de las estandarizaciones en las regiones de México para cada año del periodo 2000 a 2020 (Cuadro 1). En todas las estandarizaciones transversales, en todos los años del periodo, el norte reporta el menor rezago seguido por el centro y el sur, en este orden.

En la estandarización clásica, la jerarquía regional es la misma para todos los años y sus rangos son desiguales. En las estandarizaciones mín-máx, mín-máx restringida y la balanceada las regiones tienen el mismo rango, pero en las dos primeras la jerarquía es la misma independientemente de que los valores se incrementen o disminuyan de un año a otro. La estandarización balanceada no permite comparaciones porque todo es igual de un año a otro: Los valores en los extremos (70 y 130), media (100), stdv (24.5) y rango (60) son exactamente los mismos. Estos resultados confirman las principales limitaciones de la estandarización transversal para realizar análisis longitudinal. Es necesario aplicar las estandarizaciones longitudinales para que las comparaciones sean adecuadas, como lo señalan Mazziotta y Pareto (2022).

En el nivel de los estados, las estandarizaciones proporcionan mayor detalle geográfico de los procedimientos analizados y generan la información necesaria para la segunda tarea del análisis transversal. El estudio transversal del 2020 confirma que la jerarquía es la misma para todas las entidades federativas, independientemente de la estandarización utilizada. La investigación no proporciona información transversal en los estados para el resto del periodo para evitar redundancia ya que los resultados son similares a los obtenidos al nivel de las regiones.⁵

La segunda tarea transversal, con la información de los estados de México para el 2020, elabora diagramas de caja que combina con estadística básica para evaluar el efecto de cada estandarización en el rango, los valores máximo y mínimo y la asimetría en las tres variables del rezago educativo. *Los procedimientos de estandarización, con excepción de z_{NR} y z_{EB} , no modifican la asimetría de los datos originales.* En las estandarizaciones con amplitud de 70 a 130, solo la mín-máx restringida (z_{MP}) no tiene valores exactos en los extremos. Los resultados muestran que *un igual rango no es garantía de una mejor estandarización.* Es necesario considerar tam-

⁴ Además del rezago educativo, el rezago social incluye variables sobre el acceso a los servicios de salud, la calidad y espacios de la vivienda, los servicios básicos en la vivienda, y los bienes del hogar (CONEVAL, 2007).

⁵ La información de las estandarizaciones transversal y longitudinal está disponible en el siguiente enlace: https://github.com/jtrevino41/Pobreza-en-Mx/raw/main/Reza-go_educativo_Anexo.docx

Cuadro 1. Regiones de México. Estandarizaciones transversales, 2000 al 2020 (valores xi en porcentaje)

Región	x1					x2					x3				
	2000	2005	2010	2015	2020	2000	2005	2010	2015	2020	2000	2005	2010	2015	2020
Norte	5.0	4.4	3.5	2.7	2.4	7.0	4.4	4.0	3.1	5.3	49.0	42.8	36.8	30.6	25.0
Centro	9.6	8.4	6.7	5.2	4.4	7.7	5.1	4.5	3.4	5.8	53.1	46.1	41.2	35.1	29.1
Sur	15.3	13.8	11.4	9.3	8.3	9.5	5.8	5.5	4.0	6.4	61.0	53.2	47.8	41.8	36.0
Media	10.0	8.8	7.2	5.8	5.0	8.1	5.1	4.7	3.5	5.8	54.3	47.4	41.9	35.8	30.0
DesStd	4.2	3.9	3.3	2.7	2.4	1.1	0.6	0.6	0.4	0.5	5.0	4.4	4.5	4.6	4.5
Máx	15.3	13.8	11.4	9.3	8.3	9.5	5.8	5.5	4.0	6.4	61.0	53.2	47.8	41.8	36.0
Min	5.0	4.4	3.5	2.7	2.4	7.0	4.4	4.0	3.1	5.3	49.0	42.8	36.8	30.6	25.0
Mediana	9.6	8.4	6.7	5.2	4.4	7.7	5.1	4.5	3.4	5.8	53.1	46.1	41.2	35.1	29.1

Estandarización clásica (z)

	z1					z2					z3				
	Norte	-1.18	-1.16	-1.14	-1.11	-1.07	-1.00	-1.21	-1.04	-1.02	-1.15	-1.08	-1.05	-1.13	-1.13
Centro	-0.08	-0.11	-0.16	-0.20	-0.26	-0.36	-0.03	-0.30	-0.33	-0.13	-0.25	-0.29	-0.17	-0.16	-0.22
Sur	1.26	1.28	1.30	1.31	1.33	1.36	1.24	1.35	1.36	1.29	1.33	1.34	1.30	1.30	1.32
Media	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DesStd	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
Rango	2.45	2.44	2.43	2.43	2.41	2.37	2.45	2.39	2.38	2.44	2.41	2.40	2.43	2.43	2.42

Estandarización Mín-Máx (z_{MM})

	z1 _{MM}					z2 _{MM}					z3 _{MM}				
	Norte	70.0	70.0	70.0	70.0	70.0	70.0	70.0	70.0	70.0	70.0	70.0	70.0	70.0	70.0
Centro	97.1	95.8	94.2	92.7	90.2	86.4	98.7	88.6	87.4	95.1	90.6	89.1	93.8	93.9	91.9
Sur	130.0	130.0	130.0	130.0	130.0	130.0	130.0	130.0	130.0	130.0	130.0	130.0	130.0	130.0	130.0
Media	99.0	98.6	98.1	97.6	96.7	95.5	99.6	96.2	95.8	98.4	96.9	96.4	97.9	98.0	97.3
DesStd	24.5	24.6	24.6	24.7	24.9	25.3	24.5	25.1	25.2	24.6	24.9	25.0	24.7	24.7	24.8
Rango	60.0	60.0	60.0	60.0	60.0	60.0	60.0	60.0	60.0	60.0	60.0	60.0	60.0	60.0	60.0

Estandarización Mín-Máx restringida (z_{MP})

	z1 _{MP}					z2 _{MP}					z3 _{MP}				
	Norte	41.0	41.4	41.9	42.4	43.3	44.5	40.4	43.8	44.2	41.6	43.1	43.6	42.1	42.0
Centro	68.0	67.2	66.1	65.1	63.5	60.9	69.2	62.4	61.6	66.7	63.7	62.8	65.9	66.0	64.6
Sur	101.0	101.4	101.9	102.4	103.3	104.5	100.4	103.8	104.2	101.6	103.1	103.6	102.1	102.0	102.7
Media	70.0	70.0	70.0	70.0	70.0	70.0	70.0	70.0	70.0	70.0	70.0	70.0	70.0	70.0	70.0
DesStd	24.5	24.6	24.6	24.7	24.9	25.3	24.5	25.1	25.2	24.6	24.9	25.0	24.7	24.7	24.8
Rango	60.0	60.0	60.0	60.0	60.0	60.0	60.0	60.0	60.0	60.0	60.0	60.0	60.0	60.0	60.0

Estandarización balanceada (z_{EB})

	z1 _{EB}					z2 _{EB}					z3 _{EB}				
	Norte	70	70	70	70	70	70	70	70	70	70	70	70	70	70
Centro	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
Sur	130	130	130	130	130	130	130	130	130	130	130	130	130	130	130
Media	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
DesStd	24.5	24.5	24.5	24.5	24.5	24.5	24.5	24.5	24.5	24.5	24.5	24.5	24.5	24.5	24.5
Rango	60.0	60.0	60.0	60.0	60.0	60.0	60.0	60.0	60.0	60.0	60.0	60.0	60.0	60.0	60.0

Fuente: elaboración del autor.

bién el empate del valor máximo y mínimo entre las variables y la asimetría en cada una de ellas. Solo la estandarización balanceada (z_{EB}) logra este triple propósito y confirma que la corrección de la asimetría también ajusta a la curtosis (al corregir una se corrige la otra) (Figura 1).

El estudio longitudinal ilustra las estandarizaciones al nivel de las regiones y los estados. En las regiones, en todas las estandarizaciones y todas las variables del rezago educativo, el máximo corresponde a la región Sur en el año 2000. El mínimo corresponde al norte en el 2020, con excepción

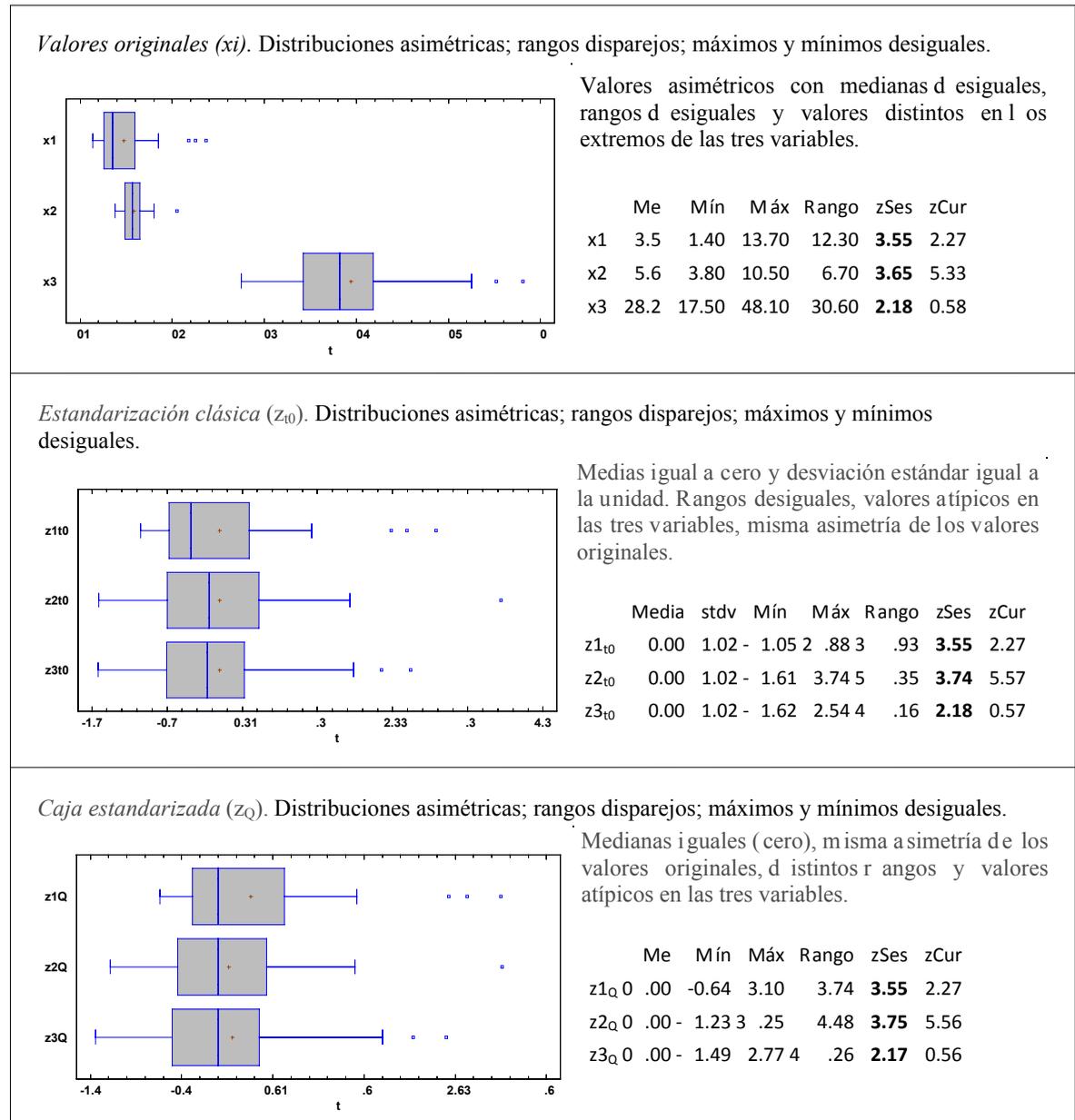
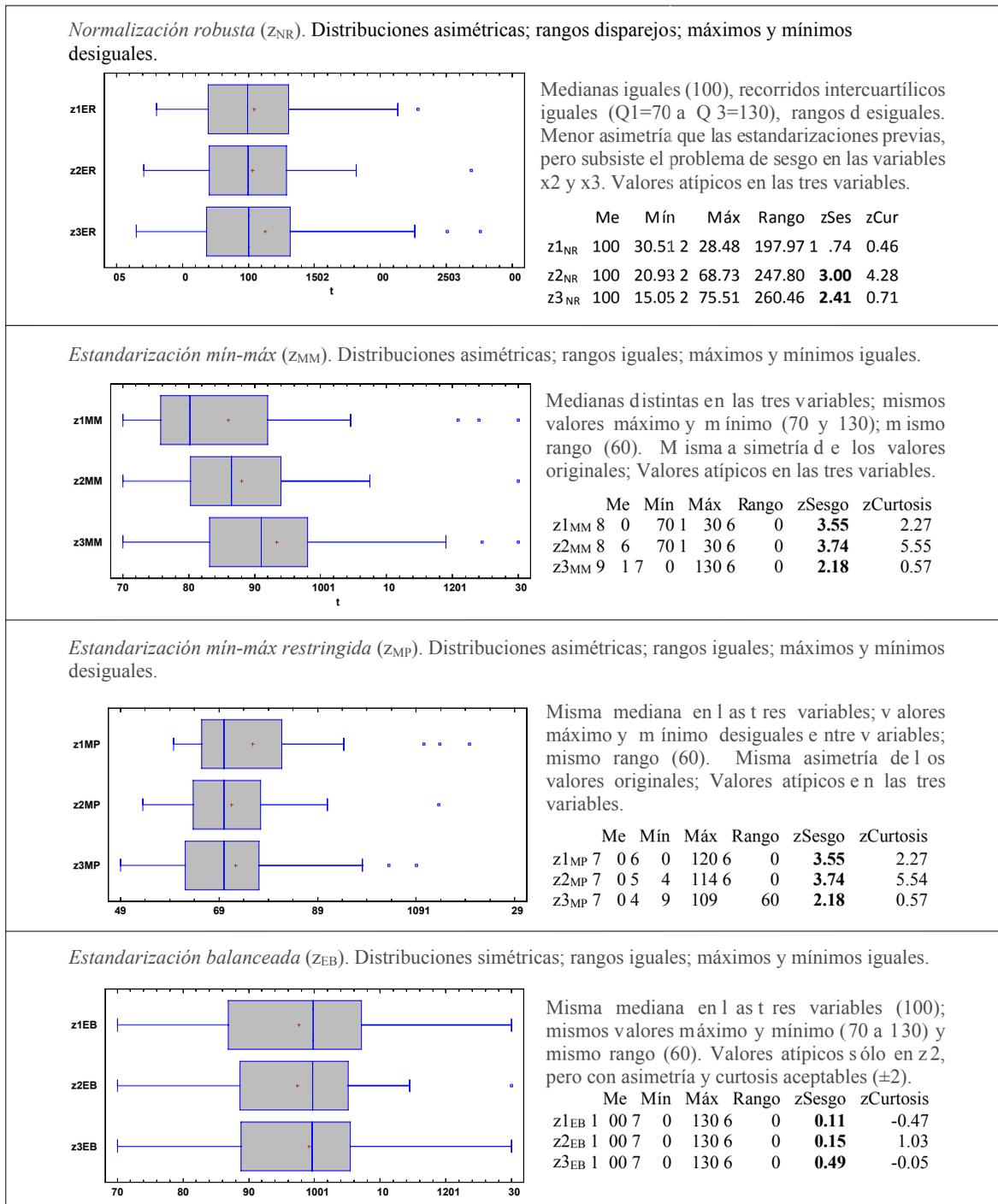


Figura 1. Variables del rezago educativo en los estados de México, 2020. Efecto de las distintas estandarizaciones en los valores originales (x_i). Fuente: Elaboración del autor en base a los datos disponibles en el enlace de la nota de pie 5.



Me = Mediana; Mín y Máx = Valor mínimo y máximo, respectivamente; Rango = Máx - Mín; zSesgo = Sesgo/Error del sesgo; zCurtosis = Curtosis/Error de la Curtosis. Los subíndices en los valores z indican lo siguiente: MM: mín-máx; MP: Mazziotta-Pareto o estandarización MM restringida; EB: estandarización balanceada. El número en z indica la variable de rezago social estandarizada. Los diagramas de caja y valores del sesgo y curtosis estandarizados son obtenidos con Statgraphics, v. 16.1.03.

Figura 1. Continuación.

de la variable z2 cuyo mínimo corresponde al 2015 (Cuadro 2). Estos resultados muestran una tendencia decreciente de las variables z1 y z3 en todas las estandarizaciones. La excepción es z2 cuyo mínimo lo alcanza el norte en el 2015. La media aritmética regional muestra claramente estas tendencias en todas las estandarizaciones. En z1 y z3 el promedio regional disminuye continuamente

para todos los años del periodo, mientras en z2 la tendencia se revierte del 2015 al 2020. El promedio de los datos originales es decreciente en todo el periodo sin mostrar el cambio en la dirección de z2 detectado por las estandarizaciones longitudinales. El estudio incluye con fines comparativos una versión de la estandarización balanceada que utiliza como referencia el promedio nacional en el 2010

Cuadro 2. Regiones de México. Estandarizaciones longitudinales, 2000 al 2020 (valores xi en porcentaje).

Año	Datos originales (x _i)					Estandarización clásica al año base (z ₀)					Estandarización Min-Máx (z _{MM})							
	Región					Media= 0; DesStd= 1, en t ₀					de 70 a 130; Ref= Me para 2000 al 2020							
	Norte	Centro	Sur	Media	DesStd	Norte	Centro	Sur	Media	DesStd	Norte	Centro	Sur	Media	DesStd			
2000	x1	5.0	7.0	49.0	20.34	20.25	z1 ₀	-1.18	-0.08	1.26	0.00	1.00	z1 _{MM}	82.2	103.8	130.0	105.33	19.54
2005		4.4	4.4	42.8	17.19	18.10		-1.34	-0.38	0.90	-0.27	0.92		79.1	97.9	123.0	100.00	17.99
2010		3.5	4.0	36.8	14.78	15.56		-1.55	-0.78	0.35	-0.66	0.78		75.1	90.0	112.2	92.41	15.24
2015		2.7	3.1	30.6	12.15	13.08		-1.73	-1.13	-0.15	-1.01	0.65		71.5	83.2	102.3	85.68	12.71
2020		2.4	5.3	25.0	10.92	10.05		-1.81	-1.33	-0.40	-1.18	0.58		70.0	79.3	97.5	82.25	11.42
2000	x2	9.6	7.7	53.1	23.48	20.94	z2 ₀	-1.00	-0.36	1.36	0.00	1.00	z2 _{MM}	106.8	113.1	130.0	116.63	9.80
2005		8.4	5.1	46.1	19.87	18.61		-3.48	-2.85	-2.16	-2.83	0.54		82.6	88.7	95.5	88.92	5.27
2010		6.7	4.5	41.2	17.45	16.80		-3.84	-3.43	-2.51	-3.26	0.56		79.0	83.0	92.0	84.70	5.45
2015		5.2	3.4	35.1	14.56	14.53		-4.76	-4.50	-3.86	-4.37	0.38		70.0	72.6	78.8	73.79	3.70
2020		4.4	5.8	29.1	13.08	11.31		-2.64	-2.19	-1.57	-2.13	0.44		90.8	95.1	101.2	95.72	4.27
2000	x3	15.3	9.5	61.0	28.59	23.01	z3 ₀	-1.08	-0.25	1.33	0.00	1.00	z3 _{MM}	109.9	116.8	130.0	118.92	8.32
2005		13.8	5.8	53.2	24.27	20.73		-2.32	-1.65	-0.22	-1.40	0.87		99.6	105.2	117.1	107.30	7.27
2010		11.4	5.5	47.8	21.58	18.73		-3.52	-2.64	-1.30	-2.49	0.91		89.6	97.0	108.1	98.22	7.59
2015		9.3	4.0	41.8	18.37	16.67		-4.75	-3.86	-2.52	-3.71	0.92		79.4	86.8	97.9	88.02	7.62
2020		8.3	6.4	36.0	16.92	13.53		-5.88	-5.07	-3.68	-4.88	0.91		70.0	76.7	88.3	78.35	7.58

Año	Estandarización Min-Máx restringida (z _{MP})					Estandarización balanceada (z _{EB})					Estandarización balanceada restringida (z _{BR})							
	de 70 a 130; Ref= Me para 2000 al 2020					de 70 a 130; Ref= Me para 2000 al 2020					de 70 a 130; Ref= Me para 2000 al 2020							
	Norte	Centro	Sur	Media	DesStd	Norte	Centro	Sur	Media	DesStd	Norte	Centro	Sur	Media	DesStd			
2000	z1 _{MP}	62.2	83.8	110.0	85.33	19.54	z1 _{EB}	88.3	110.3	130.0	109.56	17.02	z1 _{BR}	87.6	109.9	130.0	109.17	17.32
2005		59.1	77.9	103.0	80.00	17.99		83.6	106.0	124.7	104.77	16.81		83.1	105.4	124.6	104.38	16.98
2010		55.1	70.0	92.2	72.41	15.24		77.6	100.0	116.6	98.07	15.99		77.3	98.8	116.3	97.47	15.97
2015		51.5	63.2	82.3	65.68	12.71		72.3	89.8	109.3	90.43	15.10		72.2	89.0	108.8	89.99	14.97
2020		50.0	59.3	77.5	62.25	11.42		70.0	83.9	105.6	86.51	14.66		70.0	83.3	105.1	86.15	14.47
2000	z2 _{MP}	86.0	92.3	109.2	95.83	9.80	z2 _{EB}	112.2	117.1	130.0	119.77	7.50	z2 _{BR}	114.3	118.5	130.0	120.93	6.65
2005		61.8	67.9	74.7	68.12	5.27		88.1	97.0	103.6	96.24	6.34		93.9	102.0	106.6	100.84	5.24
2010		58.2	62.2	71.2	63.90	5.45		83.0	88.8	101.0	90.92	7.48		87.2	94.8	104.3	95.42	7.00
2015		49.2	51.8	58.0	52.99	3.70		70.0	73.7	82.7	75.47	5.34		70.0	74.9	86.8	77.22	7.05
2020		70.0	74.4	80.4	74.92	4.27		100.0	103.3	108.0	103.77	3.27		103.4	106.4	110.5	106.76	2.90
2000	z3 _{MP}	82.0	88.9	102.1	91.01	8.32	z3 _{EB}	111.2	117.7	130.0	119.64	7.78	z3 _{BR}	111.8	118.1	130.0	119.98	7.53
2005		71.7	77.3	89.1	79.38	7.27		101.6	106.8	117.9	108.77	6.80		102.5	107.6	118.3	109.46	6.58
2010		61.7	69.0	80.2	70.30	7.59		91.1	99.0	109.5	99.85	7.55		91.9	100.1	110.2	100.73	7.46
2015		51.5	58.9	70.0	60.10	7.62		80.1	88.0	100.0	89.36	8.19		80.5	88.7	101.0	90.06	8.42
2020		42.1	48.8	60.4	50.43	7.58		70.0	77.2	89.7	78.97	8.14		70.0	77.5	90.5	79.33	8.47

Fuente. Elaboración del autor en base a la metodología. Los números en negritas indican el valor máximo y mínimo de cada variable por periodo. Cálculos en la selección de archivos xls en https://github.com/jtrevino41/Pobreza-en-Mx/raw/main/Files_w_implicita.xlsx

(zBR), como lo sugiere la estandarización mín-máx restringida (zMP). Los resultados no agregan información relevante adicional a la obtenida por la estandarización balanceada con la mediana del periodo como referencia.⁶

La estandarización balanceada modifica los valores originales de forma consistente con el resto de las estandarizaciones. Además, como lo muestra el análisis transversal, la estandarización balanceada neutraliza el peso implícito asociado a los rangos desiguales y los valores máximo y mínimo distintos en los extremos entre las variables, y a la asimetría de cada variable. Estas características revelan que zEB es un procedimiento atractivo para estandarizar variables de forma transversal o longitudinal, especialmente las que son agregadas en un índice compuesto espacio temporal. Este argumento es suficiente para tomar como base zEB para describir la trayectoria longitudinal de las variables en los estados.

La desagregación por estados devela movimientos o cambios de dirección no detectados en el nivel regional o agregado⁷. En los estados, la ausencia de valores negativos en las diferencias en dos de las tres variables con estandarización balanceada, la Población de 15 años o más analfabeta (z1EB) y la Población de 15 años y más con educación básica incompleta (z3EB), muestra una evolución decreciente en cada uno de los quinquenios y en el periodo 2000 a 2020. La variable Población de 6 a 14 años que no asiste a la escuela (z2EB), en cambio, reporta un incremento en el quinquenio 2005 a 2010 para los estados de Colima, Cd de Mx, Morelos, Quintana Roo y Sonora, no detectado en el análisis regional. La tendencia regresa al cauce decreciente en todos los estados en el quinquenio

⁶ La posible atracción de la z_{BR} es que la referencia se mantiene fija y evita recalcular los índices al agregar nueva información. Esta ventaja no es muy convincente porque es necesario identificar el nuevo valor mínimo y máximo. Si estos valores cambian, el recálculo de valores previos es inevitable, aunque poco laborioso con cualquier equipo de cómputo actual. La supuesta ventaja de la z_{BR} desaparece porque este último argumento también es válido para el resto de los procedimientos.

⁷ La información para los estados está disponible en el enlace de la nota de pie 6.

2010 a 2015. La dirección decreciente de z2EB se detiene en el último quinquenio del periodo para dar lugar a un incremento generalizado de esta variable en todos los estados. Es probable que el incremento de z2EB del 2015 al 2020 se deba al confinamiento sanitario ocasionado por la pandemia del covid-19. Esta misma explicación aplica a los valores negativos de z2EB en el periodo 2000 a 2020 para la Cd de Mx y NL, donde la política sanitaria operó de forma inmediata.

Todos los cambios positivos o negativos, quinquenales o de todo el periodo, son estadísticamente significativos. En todos los quinquenios el año previo tiene una media superior al año sucesivo, excepto en z2 en el lapso del 2015 al 2020 (Cuadro 3). Las pruebas t para datos relacionados muestran que el declive en dos de las tres variables del reza-go educativo es estadísticamente significativo en todo el periodo y en cada uno de sus quinquenios (Cuadro 3). Los cambios en z2 también son significativos, con la particularidad de que la significancia estadística refiere un cambio de tendencia en el último quinquenio del periodo. Sin embargo, el renglón para el periodo 2000 a 2020 muestra que la tendencia general de z2 es decreciente.

Estos resultados muestran que la estandarización balanceada expresa las variables en una misma medición abstracta, controla las características que generan un peso implícito al agregarlas en un índice compuesto y permite el análisis estadístico de la tendencia espacio temporal.

DISCUSIÓN

Todas las estandarizaciones en esta investigación son adaptaciones de la estandarización clásica (z_{t0}). En las cajas estandarizadas (z_Q) la mediana reemplaza a la media y el recorrido intercuartílico a la desviación estándar. En la estandarización mín-máx (z_{MM}), el valor mínimo sustituye a la media y el rango máx-mín a la desviación estándar. La estandarización mín-máx restringida (z_{MP}) es una versión de la estandarización mín-máx donde un referente (media, mediana o una observación destacada) toma el lugar de la media y el rango máx-mín sustituye a la desviación estándar. La normalización

Cuadro 3. Pruebas t para muestras emparejadas con valores de estandarización balanceada longitudinal para las variables del rezago educativo, 2000 a 2020 (n=32)

Años	Analfabeta (z1 _{EB})				No Asiste (z2 _{EB})				Incompleta (z3 _{EB})			
	Media	Stdv	t	Sig.	Media	Stdv	t	Sig.	Media	Stdv	t	Sig.
2000	104.89	12.08	14.99	0.00	107.92	7.38	21.88	0.00	112.73	9.06	34.56	0.00
2005	101.98	12.46			96.91	7.73			105.88	8.76		
2005	101.98	12.46	16.40	0.00	96.91	7.73	4.78	0.00	105.88	8.76	25.64	0.00
2010	97.42	12.68			93.70	8.19			100.04	9.08		
2010	97.42	12.68	15.64	0.00	93.70	8.19	18.74	0.00	100.04	9.08	29.28	0.00
2015	92.18	13.25			83.85	8.86			92.74	9.31		
2015	92.18	13.25	9.82	0.00	83.85	8.86	-17.78	0.00	92.74	9.31	25.74	0.00
2020	89.22	13.16			101.65	4.85			85.45	9.41		
2000	104.89	12.08	24.11	0.00	107.92	7.38	6.14	0.00	112.73	9.06	39.89	0.00
2020	89.22	13.16			101.65	4.85			85.45	9.41		

Fuente: elaboración del autor con el programa SPSS v.26.0.0.0. Las negritas indican la inversión de la tendencia decreciente en z2. El valor negativo de t registra este cambio de tendencia.

robusta (z_{NR}) toma en cuenta los datos por encima o debajo de la mediana (Me). En ambos casos, el numerador es el mismo: la mediana toma el lugar de la media, (x_i - Me). En datos por debajo de la mediana, el término (Me - Q1) toma el lugar de la desviación estándar; en datos por encima de la mediana, (Q3 - Me) reemplaza a la desviación estándar. La estandarización balanceada (z_{EB}) es similar a la z_{NR}, solo que los valores mínimo y máximo toman el lugar del primer y tercer cuartil, respectivamente. En z_{EB} el numerador es el mismo para valores por

encima o debajo de la mediana: (x_i - Me); en el denominador, (Me - Min) o (Máx - Me) toma el lugar de la desviación estándar.

Los resultados muestran que *solo la estandarización balanceada propuesta en esta investigación corrige simultáneamente los problemas de amplitud desigual, valores mínimo y máximo dispares y asimetría que implican una ponderación implícita en la agregación de variables* (Cuadro 4).

La versión longitudinal de cada estandarización requiere que los parámetros sean fijos en el tiempo.

Cuadro 4. Estandarización. Tres efectos deseables para eliminar el peso implícito de las variables en un Índice Compuesto (IC)

Procedimiento de estandarización	Valores iguales en los extremos entre variables	Igual rango entre variables	Corrige sesgo y curtosis de cada variable
Clásica (z _{t0})	X	X	X
Cajas estandarizadas (z _Q)	X	X	X
Mín-Máx (z _{MM})	✓	✓	X
Mín-Máx restringida (z _{MP})	X	✓	X
Robusta (z _{NR})	X ^a	X ^b	✓ ^c
Balanceada (z _{EB})	✓	✓	✓
Balanceada restringida (z _{BR})	✓	✓	✓

^a Valores iguales en Q1 y Q3 de todas las variables; ^b igual recorrido intercuartílico (Q1 a Q3); ^c sólo en algunas variables. La base metodológica para los cuartiles Q1 y Q3 en z_{NR} es llevada a los valores mínimo y máximo en z_{EB} y z_{ER}.

Fuente: resumen de la información en la Figura 1.

Esto es posible si el procedimiento se aplica a datos apilados, como lo ilustra el estudio de caso.

CONCLUSIÓN

Sin estandarización, transversal o longitudinal, no hay índices compuestos confiables. La selección del procedimiento de estandarización apropiado es un reto actual y un debate en curso. El análisis comparativo de procedimientos muestra que la clave en la extensión longitudinal de las versiones transversales es mantener parámetros fijos para datos apilados, de la siguiente manera:

a) Media y desviación estándar (stdv) de un año específico fijos. La media y stdv de un año de referencia (por ejemplo, 2000) son las mismas para todos y cada de los valores originales del periodo (por ejemplo, 2000 a 2020), como en z_{t0} . En la versión de Norman (2010 y 2015) la media y desviación estándar son fijos porque corresponden a todos los valores del periodo.

b) Valor mínimo (mín) y máximo (máx) del periodo fijos, como en z_{MM} . Esto equivale a sustituir en z_{t0} la media por el valor mínimo del periodo y la desviación estándar por el rango mín-máx del periodo.

c) Media de un año específico (por ejemplo, 2010) y valores mín y máx del periodo fijos, como en z_{MP} .

d) Mediana (Me) del periodo y valor mín y máx fijos, como en como en z_{EB} .

En las variables de un índice compuesto no basta que la estandarización sea longitudinal o absoluta, es necesario también que controle la ponderación implícita al agregar variables. Los valores máximo y mínimo distintos entre variables, los rangos desiguales entre ellas y la asimetría de cada variable causan la ponderación implícita. La z_{EB} es el único procedimiento que corrige simultáneamente este problema triple.

La ponderación implícita magnifica el efecto sustitución o compensatorio en la agregación lineal de variables. La medición del efecto de la agregación de variables con y sin ponderación implícita es materia de otra fase en la elaboración de IC: la medición de la estabilidad de un IC. En la fase de

estabilidad e importancia de las variables, el efecto de procedimientos distintos de estandarización se evalúa bajo supuestos fijos de ponderación (por ejemplo, igual ponderación, entropía o juicio de expertos), con variables agrupadas o no agrupadas en dimensiones, agregadas por procedimientos compensatorios (por ejemplo, media aritmética o análisis por componentes principales) o no compensatorios (por ejemplo, media geométrica o media contra armónica y sus componentes).

El estudio propone e ilustra un procedimiento de estandarización; no sugiere ningún procedimiento de ponderación explícita ni de agregación de variables. La medición del efecto de la estandarización balanceada en un IC está fuera del alcance de este trabajo. Esta medición es materia de un escrito que retome sintéticamente las tres fases básicas de un IC: la estandarización, ponderación y agregación de variables. Cada una de estas fases es un tema de investigación por sí misma, como lo muestran el desarrollo y los resultados de esta investigación sobre estandarización.

En suma, la estandarización balanceada tiene varias funciones en un índice compuesto (IC): a) expresar los valores en una misma unidad abstracta que garantice la comparación y agregación transversal y longitudinal de variables; b) identificar casos atípicos; c) invertir la polarización de variables de manera que un incremento en la variable se refleje en un incremento en el IC; d) permitir el análisis estadístico y la prueba de hipótesis de información compatible en el tiempo y espacio; y d) controlar la ponderación implícita en: i) los rangos desiguales entre las variables, ii) los valores desiguales en los extremos entre variables, y iii) la asimetría de cada variable.

REFERENCIAS

- Angoa, I., Pérez-Mendoza, S. y Polèse, M. (2009). Los tres Méxicos: Análisis de la distribución espacial del empleo en la industria y los servicios superiores, por tamaño urbano y por región. *EURE (Santiago)*, 35(104). <https://doi.org/10.4067/s0250-71612009000100006>
- Acharya, R. y Porwal, A. (2020). A vulnerability index for the management of and response to the COVID-19

- epidemic in India: An ecological study. *The Lancet Global Health*, 8(9), e1142-e1151. Balcerzak, A. P. (2016). Multiple-criteria Evaluation of Quality of Human Capital in the European Union Countries, *Economics and Sociology*, 9(2), 11-26. <https://doi.org/10.14254/2071-789X.2016/9-2/1>
- Brimicombe, A. (2000). Constructing and evaluating contextual indices using GIS: A case of primary school performance tables. *Environment and Planning A: Economy and Space*, 32(11), 1909-1933. <https://doi.org/10.1068/a3316>
- CONEVAL. Consejo Nacional de Evaluación de la Política de Desarrollo Social. (2021). *Índice de rezago social 2020 a nivel nacional, estatal, municipal y localidades*. https://www.coneval.org.mx/Medicion/IRS/Paginas/Indice_Rezago_Social_2020.aspx
- CONEVAL. Consejo Nacional de Evaluación de la Política de Desarrollo Social. (2007). *Los mapas de pobreza en México. Anexo técnico metodológico*. México.
- Cuttillo, A., Mazziotta, M. y Pareto, A. (2021). *A composite index of economic well-being for the european union countries*. En C. Perna, N. Salvati, y F. Schirripa (Eds.), *Conference of the società italiana di statistica (SIS) 2021* (pp. 546-551). Pearson. <https://it.pearson.com/content/dam/region-core/italy/pearson-italy/pdf/Docenti/Universit%C3%A0/pearson-sis-book-2021-parte-1.pdf>
- Dasgupta, P. y Weale, M. (1992). On measuring the quality of life. *World Development*, 20(1), 119-131. [https://doi.org/10.1016/0305-750x\(92\)90141-h](https://doi.org/10.1016/0305-750x(92)90141-h)
- Exeter, D. J., Boyle, P. J. y Norman, P. (2011). Deprivation (im)mobility and cause-specific premature mortality in Scotland. *Social Science & Medicine*, 72(3), 389-397. <https://doi.org/10.1016/j.socscimed.2010.10.009>
- Flanagan, B. E., Gregory, E. W., Hallisey, E. J., Heitgerd, J. L. y Lewis, B. (2011). A social vulnerability index for disaster management. *Journal of Homeland Security and Emergency Management*, 8(1). <https://doi.org/10.2202/1547-7355.1792>
- Gilthorpe, M. S. (1995). The importance of normalization in the construction of deprivation indices. *Journal of Epidemiology & Community Health*, 49(Suppl 2), S45-S50. https://doi.org/10.1136/jech.49.suppl_2.s45
- Heinrich, K., Huber, C. y Rinner, C. (2017). Making Area-based Composite Indices Comparable across Time—The Role of Standardization and Outlier Removal. *URISA Journal*, 27(2), 37-49. <https://www.urisa.org/clientuploads/directory/Documents/Journal/Vol27No2.pdf>
- Mazziotta, M. y Pareto, A. (2021). Everything you always wanted to know about normalization (but were afraid to ask). *Rivista Italiana di Economia Demografia e Statistica*, 75(1), 41-52. http://www.sieds.it/listing/RePEc/journl/2021751P041_052_Mazziotta.pdf
- Mazziotta, M. y Pareto, A. (2022). Normalization methods for spatio-temporal analysis of environmental performance: Revisiting the Min–Max method. *Environmetrics*, 33(5). <https://doi.org/10.1002/env.2730>
- Norman, P. (2010). Identifying change over time in small area socio-economic deprivation. *Applied Spatial Analysis and Policy*, 3(2-3), 107-138. <https://doi.org/10.1007/s12061-009-9036-6>
- Norman, P. (2015). The changing geography of deprivation in Britain: exploiting small area census data 1971 to 2011. En N. Malleon, N. Addis, H. Durham, A. Heppenstall, R. Lovelace, P. Norman, y R. Oldroyd (Eds.), *GIS Research UK (GISRUK) 2015 Proceedings* (pp. 465-474). School of Geography, University of Leeds. https://figshare.com/articles/journal_contribution/GIS_Research_UK_GISRUK_2015_Proceedings/1491375
- OECD. (2008). *Handbook on constructing composite indicators. Methodology and user guide*. Paris: OECD Publications. <https://www.oecd.org/sdd/42495745.pdf>
- Ricketts Hein, J., Ilbery, B. y Kneafsey, M. (2006). Distribution of local food activity in England and Wales: An index of food relocalization. *Regional Studies*, 40(3), 289-301. <https://doi.org/10.1080/00343400600631533>
- Rinner, C. y Pietropaolo, A. (2021). Tracking area-based composite indices over time: change indicators vs. jointly rescaled indicators. *Geomatica*, 75(2), 51–64. <https://cdnsiencepub.com/doi/pdf/10.1139/geomat-2020-0014>
- Sibley, D. (1987). *Spatial Applications of Exploratory Data Analysis*. Geo Books, Norwich.