Desigualdad social y desigualdad educativa

Francesc J. Hernàndez (Universitat de València), noviembre de 2021

1.-Introducción

Agradezco mucho la invitación para participar en el XXIII Encuentro Nacional de Investigación Educativa, promovido por el IMCED de México. Es un honor para mi tomar parte en este evento y una gran satisfacción poderme encontrar con personas comprometidas con la educación.

He de declarar que mi objetivo con este texto es ambicioso. Me gustaría hacer una presentación científica, en sentido enfático, sobre la relación entre desigualdad social y desigualdad educativa.

Comenzaré analizando la partícula «y» del título que vincula dos fenómenos: la desigualdad social y la desigualdad educativa, que consideraré más adelante. Una presentación científica de esa conjunción exige el cálculo de la correlación entre ambos fenómenos.

He de advertir que «correlación» no significa «causalidad». La ciencia es muy cauta a la hora de atribuir causalidades, como demuestra clásicamente la historia del doctor Ignaz Semmelweis y sus investigaciones sobre las fiebres puerperales en el Hospital General de Viena. El gran científico Richard Feynman explicaba estas precauciones a partir de los experimentos del investigador Young con las ratas, sumamente precavido al eliminar cualquier factor que pudiera inducirle a error con sus resultados (Feynman 1974). Por ello, aquí no estoy en disposición de plantear *causalidades*, sino solo de establecer *correlaciones*, o como mucho de aproximarnos a esas causalidades.

La estadística ha proporcionado diversos coeficientes de correlación, el más célebre de los cuales es el de Pearson-Bravais (para simplificar, de Pearson). A continuación explicaré su funcionamiento. Si la lectora o el lector lo conoce puede prescindir del siguiente epígrafe.

2.-El coeficiente de correlación de Pearson

El coeficiente de correlación de Pearson se suele expresar mediante la letra griega ro (ϱ), la letra minúscula erre (r) o su mayúscula (R), que es la forma más extendida a partir del uso de programas de hojas de cálculo.

El coeficiente de Pearson expresa la relación entre dos variables, x e y, y su valor puede oscilar entre -1 y 1. En caso de que haya una correlación intensa y directa entre x e *y,* el valor de R se aproxima a 1. Esto quiere decir que cuando los valores de una variable crecen, los correspondientes de la otra suelen crecer al mismo ritmo, y viceversa. En caso de que no exista prácticamente correlación, el valor de R se aproxima a 0. Entonces, la oscilación de una variable no se corresponde con la de la otra. En caso de que una variable aumente y la otra disminuya al mismo ritmo, el valor de R se acerca a -1. Hablaremos entonces de una correlación intensa, pero inversa. Sin embargo, simplemente con emplear la variable inversa de una de las dos (digamos x^{-1} o y^{-1}), R cambia su signo. Por ello, muchas veces podemos utilizar el valor absoluto de R (es decir: $|R|$), sin tener en cuenta si es positivo o negativo su valor.

El coeficiente de correlación de Pearson se define como el cociente de la *covarianza* de dos variables, x e y, por el producto de sus *desviaciones típicas*.

$$R_{xy}\frac{\sigma_{xy}}{\sigma_x\ \sigma_y}$$

Veamos estos conceptos.

La *covarianza*, σ_{xy}, se define como el sumatorio del producto de la diferencia de cada valor de x menos su media por el correspondiente valor de y menos su media, dividido por el número de valores -1.

$$\sigma_{xy} \frac{\Sigma\,(x_i - \bar{x})(y_i - \bar{y})}{n-1}$$

La *desviación típica*, que se expresa con la letra *s* o la letra griega equivalente sigma, *σ*, equivale a la raíz cuadrada de la *varianza*. La varianza es la suma de la diferencia de cada valor con la media dividida por el número de valores.

$$\sigma_{\mathrm{x}}^2 = \frac{1}{n}\Sigma\,(x_i - \bar{x})$$

Una manera sencilla de calcular el coeficiente de correlación de Pearson es mediante la fórmula siguiente:

$$R_{xy} = \frac{n\,\Sigma\,x_i\,y_i - \Sigma\,x_i\,\Sigma\,y_i}{\sqrt{n\,\Sigma\,x_i^2 - (\Sigma\,x_i)^2}\,\sqrt{n\,\Sigma\,y_i^2 - (\Sigma\,y_i)^2}}$$

Dados *n* valores de las variables *x* e *y*, para calcular manualmente un coeficiente de correlación de este modo, hay que averiguar los datos de las celdas sombreadas en el esquema y sustituirlos en la fórmula anterior.

Tabla 1. Cálculo de R

x	**y**	**xy**	**x^2**	**y^2**
x_1	y_1	x_1y_1	x_1^2	y_1^2
x_2	y_2	x_2y_2	x_2^2	y_2^2
…	…	…	…	…
x_n	y_n	x_ny_n	x_n^2	y_n^2
$\Sigma\,x_n$	$\Sigma\,y_n$	$\Sigma\,x_ny_n$	$\Sigma\,x_n^2$	$\Sigma\,y_n^2$
$\Sigma\,x_n^2$	$\Sigma\,y_n^2$			

De la fórmula anterior se deduce fácilmente que $R_{xy} = R_{yx}$.

Afortunadamente, programas informáticos habituales como las hojas de cálculo (excel, etc.) permiten el *cálculo inmediato* del coeficiente de correlación de Pearson. Incluso al trazar la línea de tendencia de una dispersión de puntos en unos ejes cartesianos se puede pedir que el programa calcule R^2 (un estadístico que se denomina «coeficiente de determinación»). Hay que tener presente que si la línea de tendencia recta es creciente $R = \sqrt[2]{R^2}$, pero si la línea de tendencia es decreciente entonces $R = -\sqrt[2]{R^2}$

Hay que advertir que R es muy sensible a valores que discrepen de la tendencia general.

3.- El coeficiente de Gini

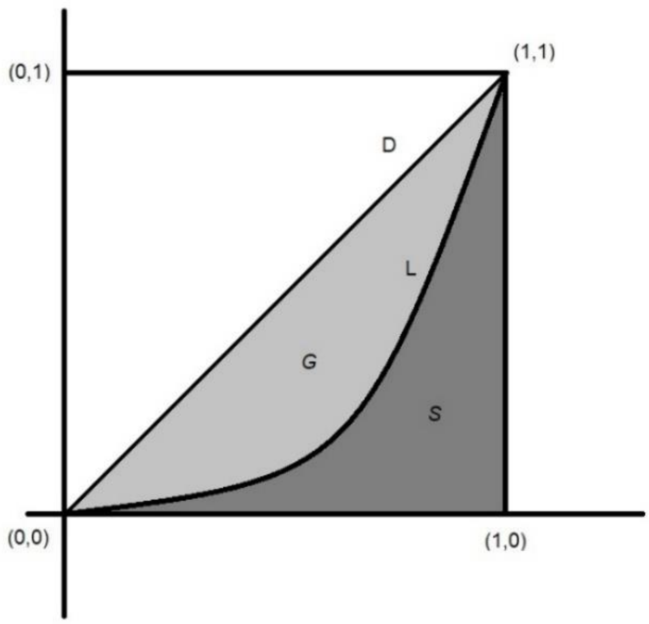
Vamos a tratar ahora de la desigualdad social. Podemos hablar largamente sobre la desigualdad social, pero en ciencias sociales es un objeto que se mide de dos maneras diferentes: mediante comparación de cuantiles o mediante el Coeficiente de Gini.

Los cuantiles son las partes de una muestra ordenada. En estadística se suelen considerar los cuartiles (cuarta parte), los quintiles (quinta parte), los deciles (décima parte; a veces en femenino: décilas) o los centiles o percentiles (centésima parte) de una muestra ordenada de menor a mayor.

Por ejemplo, la Oficina de Estadística de la Unión Europea (Eurostat) mide la desigualdad de los diversos países con una proporción entre quintiles. Este indicador se abrevia como «S80/S20» (la *S* de *share*), y se entiende que es la proporción entre el quintil más rico (lo que están por encima del 80% de la muestra) y el quintil más pobre (los que están por debajo del 20%, por eso la abreviatura) (EUROSTAT 2021). La OCDE (OECD, en inglés) mide no solo S80/S20 sino también la proporción entre el valor superior del noveno decil (esto es, el punto 90) y el límite superior del primer decil (el punto 10), que se representa P90/P10 (la P por *point*). También, considerando que el punto 50 corresponde a la renta media, se usan como indicadores P90/P50 y P50/P10. Así mismo se calcula la denominada proporción de Palma, que podríamos definir como S90/S40 (OECD 2021).

Las comparaciones de cuantiles son susceptibles de recibir, entre otras, la crítica de que comparan dos fragmentos de la muestra, pero ignoran el resto. Tal vez por ello se ha generalizado otro índice o coeficiente para medir la desigualdad social, el propuesto por el economista italiano Corado Gini y que toma su nombre: el índice o coeficiente de Gini. Se denomina índice cuando oscila entre 0 y 1, y coeficiente cuando su valor se encuentra entre 0 y 100; el coeficiente es el índice multiplicado por 100. Aquí nos referiremos al coeficiente (en la bibliografía anglófona generalmente se dice índice aunque oscile entre 0 y 100). Para entender la lógica del coeficiente de Gini, obsérvese el gráfico siguiente:

Gráfico 1. Explicación del Coeficiente de Gini con la curva de Lorenz



Si consideramos el porcentaje de población acumulada en el eje X (podemos anotarla como un porcentaje, de 0% a 100%, o con valores entre 0 y 1) y el porcentaje de ingresos acumulados en el eje Y, obtenemos una curva (L) –denominada curva de Lorenz– que está por debajo de la diagonal (D). Si la línea L coincidiera con la diagonal D estaríamos ante una igualdad ideal, lo que no es el caso. La proporción entre el área definida por la curva L y la diagonal D (marcada en el gráfico por la letra G y una zona de color gris claro) y el triángulo (0,0), (1,0) y (1,1) (el área $G + S$, la zona gris clara y la gris oscura, cuyo valor es la mitad del cuadrado unidad, por tanto $\frac{1}{2}$) se conoce como índice de Gini (recuérdese que multiplicado por 100 proporciona el coeficiente).

$$Coef.Gini = 100\frac{G}{G+S} = 100\frac{G}{\frac{1}{2}} = 200G$$

La OCDE y el Banco Mundial ofrecen datos del Coeficiente de Gini para los diversos países del mundo. Las políticas de estas instituciones pueden ser cuestionables, pero no tenemos motivos para dudar de la objetividad de estos indicadores.

A continuación se explica con más detalle su cálculo, en un nuevo epígrafe de ampliación cuya lectura puede resultar prescindible.

5.- El cálculo del coeficiente de Gini

En términos prácticos, para establecer el área de G, precisamos calcular el *área bajo la curva* L, lo que se puede hacer como una integral definida de la función de la curva entre los puntos 0 y 1 del eje X. Como puede verse en la fórmula posterior, en la que hemos utilizado una función polinómica de sexto grado, el hecho de proceder con porcentajes acumulados convierte una fórmula muy complicada (la integral definida) en una simple suma de fracciones, lo que representa una magnífica simplificación. Supongamos que tenemos una ecuación de la curva L del tipo: $ax^6 + bx^5 + cx^4 + dx^3 + ex^2 + fx + g$, entonces:

$$S = \int_0^1 (ax^6 + bx^5 + cx^4 + dx^3 + ex^2 + fx + g)dx =$$

$$a\int_0^1 x^6\,dx + b\int_0^1 x^5\,dx + c\int_0^1 x^4\,dx + d\int_0^1 x^3\,dx + e\int_0^1 x^2\,dx + f\int_0^1 x\,dx + g\int_0^1 dx =$$

$$a\left[\frac{x^7}{7}\right]_0^1 + b\left[\frac{x^6}{6}\right]_0^1 + c\left[\frac{x^5}{5}\right]_0^1 + d\left[\frac{x^4}{4}\right]_0^1 + e\left[\frac{x^3}{3}\right]_0^1 + f\left[\frac{x^2}{2}\right]_0^1 + g\,[x]_0^1 =$$

$$\frac{a}{7} + \frac{b}{6} + \frac{c}{5} + \frac{d}{4} + \frac{e}{3} + \frac{f}{2} + g$$

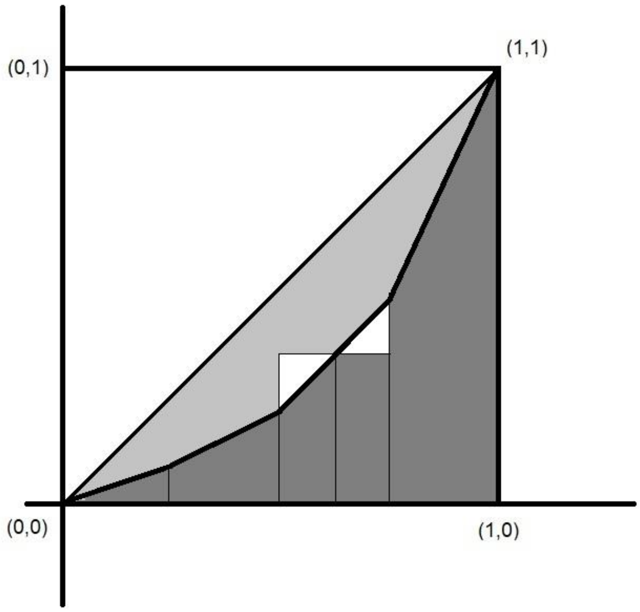
Si, por ejemplo, tuviéramos como función de la curva $ax^4 + bx^3 + cx^2 + dx + e$, entonces

$$S = \frac{a}{5} + \frac{b}{4} + \frac{c}{3} + \frac{d}{2} + e$$

Además, si ponemos pares de puntos con una hoja de cálculo en los ejes cartesianos (dentro del cuadrado unitario), podemos solicitar al programa que nos trace la línea de tendencia polinómica y nos devuelva la función (hasta grado 6). Después, los parámetros de la ecuación se pueden trasladar directamente a la suma de fracciones.

En el caso de proceder con *cuantiles*, esto es, con porciones de una muestra ordenada (como es el caso con los *cuartiles* del índice socioeconómico que comentaremos más adelante), el procedimiento es más sencillo, porque simplemente se precisa calcular el área del triángulo y los trapecios inferiores. El área de estos se puede simplificar como la semisuma de las alturas por la base, que en todos los casos es la misma (en el caso de los cuartiles: 0,25), como puede verse en el gráfico 2.

Gráfico 2. Explicación del Coeficiente de Gini con cuantiles (cuartiles)



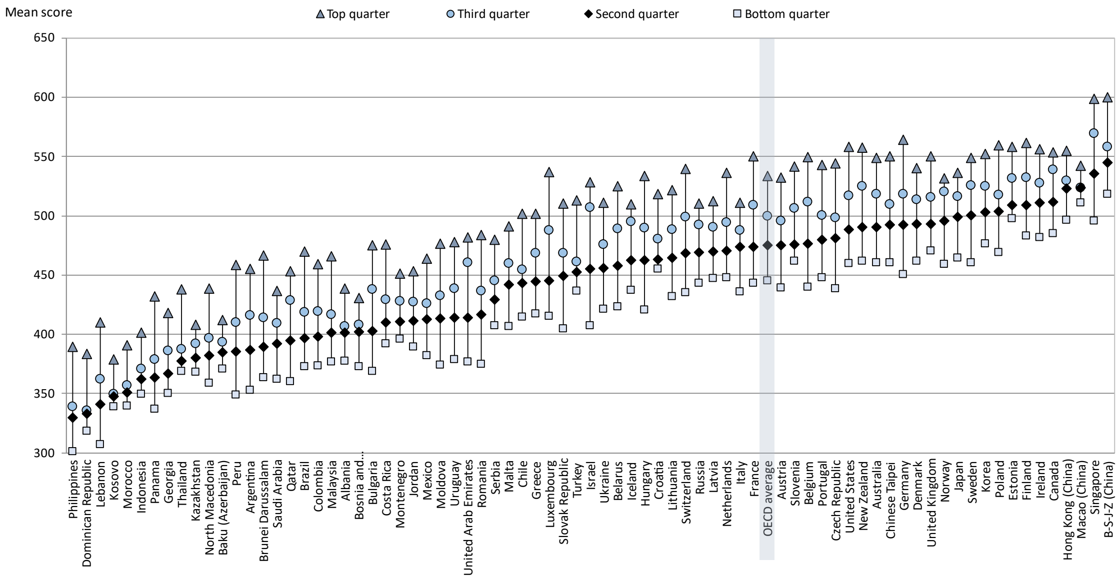
6.- La desigualdad educativa

Podríamos en este punto proceder a imaginar un experimento. Los experimentos imaginados son muy conocidos en algunos ámbitos de la ciencia, como por ejemplo la física cuántica; tal vez el más célebre sea el del gato de Schrödinger. Pero no me refiero a este tipo de experimentos irrealizables, sino a aquellos otros que se pensaron en un tiempo pero que entonces no había condiciones para poder realizarlos, y se han podido efectuar mucho tiempo después. Pues bien, plantearé la siguiente cuestión: ¿Cómo tendría que ser un experimento para medir la desigualdad educativa? En primer lugar, tendríamos que referirnos a rendimiento o logro educativo. Lógicamente el rendimiento o el logro no es el único objetivo de los sistemas educativos (por ejemplo, también lo son reducir el abandono educativo, favorecer el aprendizaje a lo largo de la vida, combatir la discriminación por razón de género, etnia u otras, etc.). Pero no hay duda de que las medidas sobre rendimiento resultan centrales en la educación y tienen una larga tradición de expresión numérica. En segundo lugar, nuestro experimento no se tendría que circunscribir a un centro educativo, a una localidad, a una región o a un país, porque intentamos conseguir resultados de validez general. En tercer lugar, si consiguiéramos una evaluación de rendimiento para muchos países esta debería versar sobre alguna materia nuclear en el currículum. Por último, tendríamos que disponer de los datos desagregados por una variable de desigualdad social, como las que hemos indicado anteriormente.

Pues bien, todas las características que podemos imaginar en un experimento para medir la desigualdad educativa ya se dan en los resultados que proporcionan las pruebas del Programme for International Student Assessment (conocido por las siglas PISA) de 2018. Como ya hemos dicho, las pruebas PISA son susceptibles de múltiples críticas, pero ello no resta validez a los resultados que consideraremos aquí.

El volumen II de los resultados de PISA (OECD 2019), presenta un significativo gráfico (figura II.2.3), en el que podemos observar cómo los diversos cuartiles de estudiantado, según el estatus socioeconómico (es decir, la clase social), presentan rendimientos correlativos en lengua. En todos los países estudiados (de los que se excluyó precisamente a España por deficiencias en las respuestas a la prueba), el cuartil inferior presenta un rendimiento menor que el segundo cuartil, este que el tercero y, a su vez, menor que el superior. Este mismo resultado también se pudo constatar en las tandas anteriores de PISA y con las otras materias investigadas por el programa. Véase el gráfico 3.

Gráfico 3. Puntuaciones en lengua según cuartiles de índice socioeconómico



Se podría pensar que este gráfico ya *demuestra* la relación entre desigualdad social y desigualdad educativa, pero no es así. Este gráfico la *muestra*, pero para establecer una demostración científica debemos *medir*, y este es un paso necesario. El objetivo de este texto es *medir* si las diferencias en los rendimientos diferenciados se correlacionan con la desigudaldad social. El mismo informe PISA parece sugerir esta relación cuando presenta en la tabla II.B1.2.1 el valor del Coeficiente de Gini para los Estados, según los cálculos del Banco Mundial para 2015, pero no profundiza en este asunto. Lo que haremos a continuación será precisamente utilizar la misma metodología del coeficiente de Gini (véase el epígrafe 5), un instrumento estadístico aceptado para medir desigualdades, aplicándolo a los rendimientos educativos y, posteriormente, establecer el coeficiente de correlación de Pearson (véase el epígrafe 2).

7.- Los datos de desigualdad educativa y social

En primer lugar necesitamos los datos del gráfico 3 y los datos del coeficiente de Gini en los diversos países. PISA proporciona los datos de 2015, por lo que se ha procedido a actualizarlos para el años de la prueba (2018) o, caso de no disponer de este, el inmediatamente anterior (2017) (The World Bank 2021a y 2021b). De este modo se forma la tabla 2.

Tabla 2. Rendimiento en lengua por cuartiles de índice socioeconómico, según PISA 2018, y Coeficiente de Gini

Reading performance, by socio-economic status (ESCS) 2018	Bottom quarter of ESCS	Second quarter of ESCS	Third quarter of ESCS	Top quarter of ESCS	Gini Index
Austria	440	475	496	533	30,8
Belgium	440	477	512	550	27,2
Chile	415	443	455	502	44,4
Colombia	373	398	419	459	50,4
Czech Republic	439	481	498	544	25,0
Denmark	462	493	514	540	28,2
Estonia	497	509	531	558	30,3
Finland	483	509	533	562	27,3
France	443	474	509	550	32,4
Germany	450	492	518	564	31,9
Greece	417	444	468	502	32,9
Hungary	420	463	489	534	29,6
Ireland	482	511	527	557	31,4
Italy	436	474	487	511	35,9
Latvia	447	470	490	512	35,1
Lithuania	432	464	488	522	35,7
Luxembourg	415	445	488	537	35,4
Netherlands	448	470	495	536	28,1
Norway	459	496	520	532	27,6
Poland	469	504	518	560	30,2
Portugal	448	480	501	543	33,5
Slovak Republic	404	449	468	511	25,0
Slovenia	462	476	506	541	24,6
Sweden	460	501	526	549	30,0
Switzerland	435	469	499	539	33,1
Turkey	437	452	461	513	41,9
United Kingdom	471	493	516	550	35,1
Belarus	423	458	489	525	25,2
Brazil	373	397	419	470	53,9
Costa Rica	392	410	429	476	48,0

Croatia	455	463	480	518	29,7
Cyprus	389	416	439	459	32,7
Dominican Republic	319	333	336	383	43,7
Georgia	350	367	386	418	36,4
Indonesia	350	362	371	402	37,8
Kazakhstan	368	380	392	408	27,8
Malaysia	377	401	417	466	41,1
Malta	406	442	460	491	28,7
Moldova	374	414	433	476	25,7
North Macedonia	359	382	397	439	33,0
Panama	337	364	379	432	49,2
Peru	349	385	410	458	42,4
Philippines	301	330	339	389	42,3
Romania	375	417	437	484	35,8
Russia	443	469	493	510	37,5
Serbia	407	429	445	480	36,2
Thailand	369	377	388	438	36,4
Ukraine	422	456	476	511	26,1
Uruguay	379	414	439	478	39,7

Con los datos de los cuartiles, podemos calcular el coeficiente de Gini de la desigualdad educativa y establecer su correlación de Pearson con el coeficiente de Gini de la desigualdad social y así demostraremos si existe o no existe este vínculo y, lo que es todavía mejor, lo mediremos.

Para explicar cómo se procede, tomemos el primer caso, Austria. Sus puntuaciones son:

Primer cuartil: 440

Segundo cuartil: 475

Tercer cuartil: 496

Cuarto cuartil: 533.

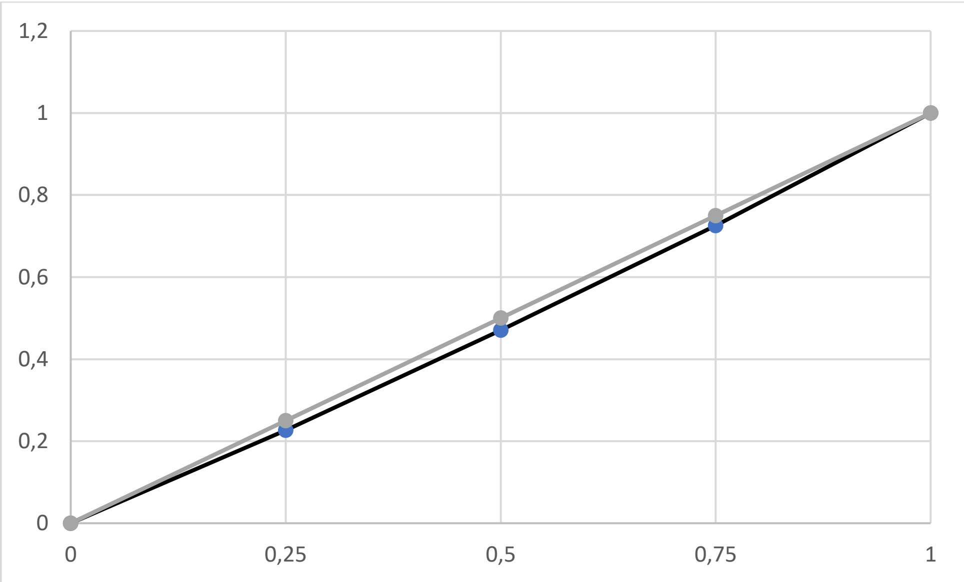
Se calcula la suma de los cuatro valores (440+475+496+533= 1944) y la proporción que representa cada uno de ellos:

Tabla 3. Ejemplo de cálculo para Austria

x	Cálculo de y	y
0,00	0	0,000
0,25	400/1.944	0,22634
0,50	(400+475)/1.944	0,47068
0,75	(400+475+496)/1.944	0,72582
1,00	(400+475+496+533)/1.944	1,000

En el gráfico siguiente se exponen los valores (x,y) unidos por una línea quebrada (de color negro) y la diagonal mayor (de color gris)

Gráfico 4. Representación gráfica del caso de Austria



Para calcular la superficie G, calcularemos el área debajo de la línea quebrada (S) por el procedimiento comentado. Será:

$$S = \left(\frac{0{,}22634}{2} * 0{,}25\right) + \left(\frac{0{,}22634 + 0{,}47068}{2} * 0{,}25\right) + \left(\frac{0{,}47068 + 0{,}72582}{2} * 0{,}25\right) + \left(\frac{0{,}72582 + 1}{2} * 0{,}25\right) = 0{,}48071$$

Por tanto, la superficie G (entre ambas líneas) és:

$$0{,}5 - 0{,}48071 = 0{,}01929$$

Y finalmente calculamos la proporción entre esta cantidad y 0,5 (que multiplicada por 100 proporciona el coeficiente):

$$Coeficiente\ de\ Gini_{desigualdad\ educativa\ Austria} = \frac{0{,}01029}{0{,}5} = 3{,}858$$

8.- El cálculo de la correlación

Antes de seguir adelante con los cálculos debemos resolver una cuestión. Las diferencias de rendimiento en puntuación PISA para lengua entre los cuartiles de ingresos es muy inferior a la de ingresos, por lo que el área resultante es comparativamente mucho más pequeña pequeña. Por ejemplo, en el primer caso, Austria, el valor del coeficiente de Gini según el Banco Mundial es 30,8, mientras que el coeficiente de Gini de las puntuaciones PISA solo es de 3,8, como hemos visto. Al establecer una correlación entre los coeficientes de Gini del Banco Mundial y los de PISA, que se mueven en una horquilla reducida, se pueden producir distorsiones del Coeficiente de Correlación de Pearson, que es muy sensible a las desviaciones como ya expliqué.

Para maximizar las diferencias entre las puntuaciones se puede recurrir a elevar cada uno de los valores de rendimiento a una potencia que denominaremos α.

Para que se entienda el procedimiento, volveremos a calcular los valores de Austria para $\alpha = 2$, el caso más sencillo.

Primer cuartil: 440; 440^2=193.600

Segundo cuartil: 475; 475^2=225.625

Tercer cuartil: 496; 496^2=246.016

Cuarto cuartil: 533; 533^2=284.089

Se calcula la suma de los cuatro valores (193.600+225.625+246.016+284.089 = 949.330) y la proporción que representa cada uno de ellos:

Tabla 4. Ejemplo de cálculo para Austria con α=2

x	Cálculo de y	y
0,00	0	0,000
0,25	193.600/949.330	0,20393
0,50	(193.600+225.625)/ 949.330	0,44160
0,75	(193.600+225.625+246.016)/ 949.330	0,70074
1,00	(193.600+225.625+246.016+284.089)/ 949.330	1,000

Para calcular la superficie G, calcularemos el área debajo de la línea quebrada (S) por el procedimiento comentado. Será:

$$S = \left(\frac{0,20393}{2} * 0,25\right) + \left(\frac{0,20393 + 0,44160}{2} * 0,25\right) + \left(\frac{0,44160 + 0,70074}{2} * 0,25\right) + \left(\frac{0,70074 + 1}{2} * 0,25\right) = 0,46156$$

Por tanto, la superficie G (entre ambas líneas) és:

$$0,5 - 0,46156 = 0,038432$$

Y finalmente calculamos la proporción entre esta cantidad y 0,5 (que multiplicada por 100 proporciona el coeficiente):

$$Coeficiente\ de\ Gini_{desigualdad\ educativa\ Austria} = \frac{0,038432}{0,5} = 7,6865$$

Como se puede ver, el valor del Coeficiente se ha duplicado. Continuaríamos así calculando todos los países y sucesivos exponentes α.

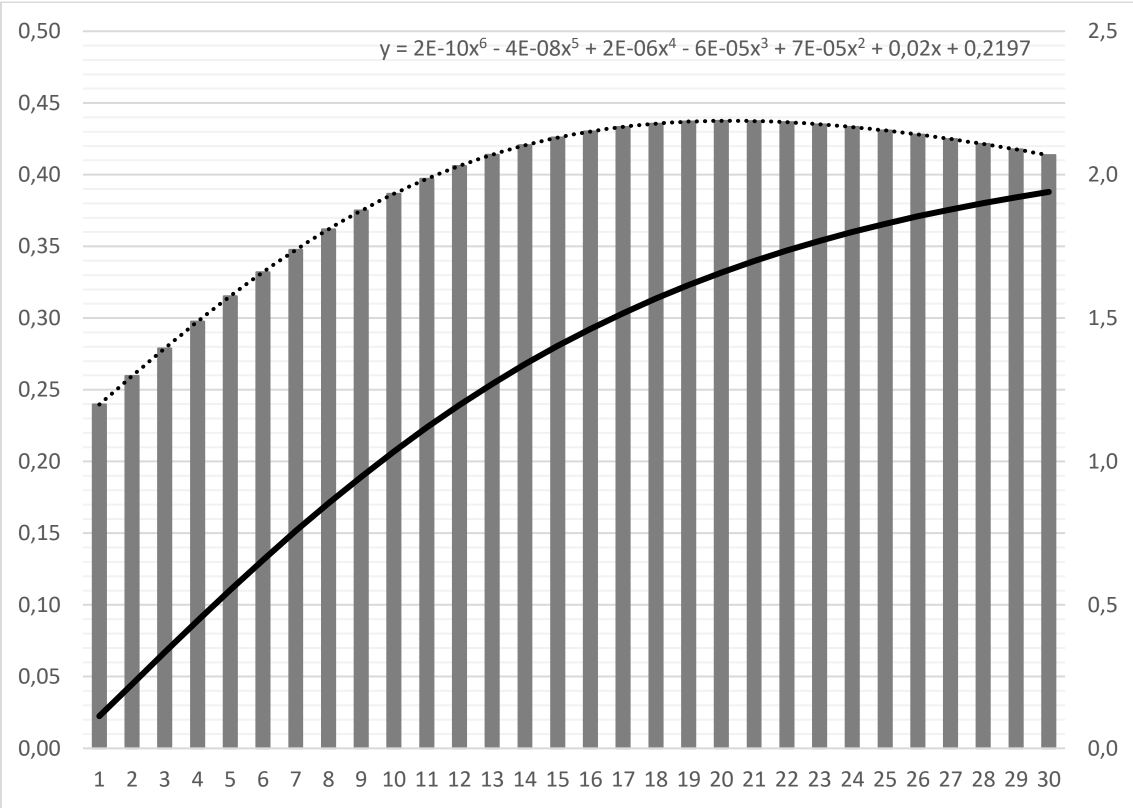
El valor del coeficiente de Correlación de Pearson entre los valores del coeficiente de Gini calculado por el Banco Mundial que mide la desigualdad social y los coeficientes de Gini sobre los datos de PISA, que miden la desigualdad educativa, variando el exponente α, se recogen en la tabla 5, donde se ha hecho constar también la media de los valores de Gini de PISA que presentan los diversos Estados y su proporción sobre la media de los valores de Gini del Banco Mundial (para los datos de la tabla anterior=34,33265).

Tabla 5. Exponentes, Coeficientes de Correlación de Pearson, Media de los valores del Coeficiente de Gini para los datos de PISA y proporción de estos respecto de la media de los valores del Coeficiente de Gini del Banco Mundial.

α	R	Media GINI PISA	Media Gini PISA/media GINI Banco Mundial (=34,33265)
1	0,23970	3,84981	0,11213
2	0,25961	7,69694	0,22419
3	0,27896	11,51520	0,33540
4	0,29757	15,27907	0,44503
5	0,31530	18,96442	0,55237
6	0,33199	22,54925	0,65679
7	0,34755	26,01432	0,75771
8	0,36189	29,34349	0,85468
9	0,37494	32,52395	0,94732
10	0,38666	35,54633	1,03535
11	0,39706	38,40449	1,11860
12	0,40614	41,09536	1,19698
13	0,41393	43,61860	1,27047
14	0,42048	45,97618	1,33914
15	0,42585	48,17204	1,40310
16	0,43011	50,21161	1,46250
17	0,43333	52,10153	1,51755
18	0,43560	53,84921	1,56845
19	0,43697	55,46263	1,61545
20	0,43755	56,95002	1,65877
21	0,43739	58,31969	1,69867
22	0,43658	59,57984	1,73537
23	0,43517	60,73848	1,76912
24	0,43324	61,80330	1,80013
25	0,43084	62,78163	1,82863
26	0,42803	63,68038	1,85481
27	0,42485	64,50603	1,87885
28	0,42136	65,26464	1,90095
29	0,41759	65,96181	1,92126
30	0,41359	66,60273	1,93992

La curva que describen los valores de R se puede apreciar más claramente en el gráfico 5, donde estos aparecen indicados como barras con la escala de la izquierda (con la línea de tendencia polinómica de grado 6 y su correspondiente ecuación), mientras que la proporción entre las medias (columna de la derecha de la tabla anterior) se representan mediante la línea de la escala de la derecha.

Gráfico 5. Correlaciones de Pearson y proporción de medias entre los Coeficientes de Gini (BM y PISA)



La tabla y la gráfica anteriores permiten aclarar un asunto más. ¿Qué valor de α será el adecuado? Podemos ver cómo las mayores aproximaciones a una proporción de las medias de los dos Coeficientes de Gini, el calculado por el Banco Mundial como indicador de desigualdad social y el calculado aquí como indicador de desigualdad educativa, se dan en el intervalo $9 < \alpha < 10$, que presenta valores de $0{,}37 < \mathrm{R} < 0{,}39$, más que satisfactorios para para $n = 49$ países, lo que podemos considerar que acredita científicamente que la desigualdad educativa correlaciona con la desigualdad social. Pero aún R eleva a 0,4375 para $\alpha = 20$.

Con ello hemos demostrado que existe una correlación entre la desigualdad social y la desigualdad educativa.

8.- Ulteriores análisis

Podemos estudiar aún más estos resultados y realizar análisis ulteriores. Aquí se ofrece un ejemplo. Es posible tomar el caso de $\alpha = 10$ (el mayor en la proximidad de la proporción $= 1$) y estudiar la variación de R en caso de prescindir de cada Estado. Los resultados se recogen en la tabla 6.

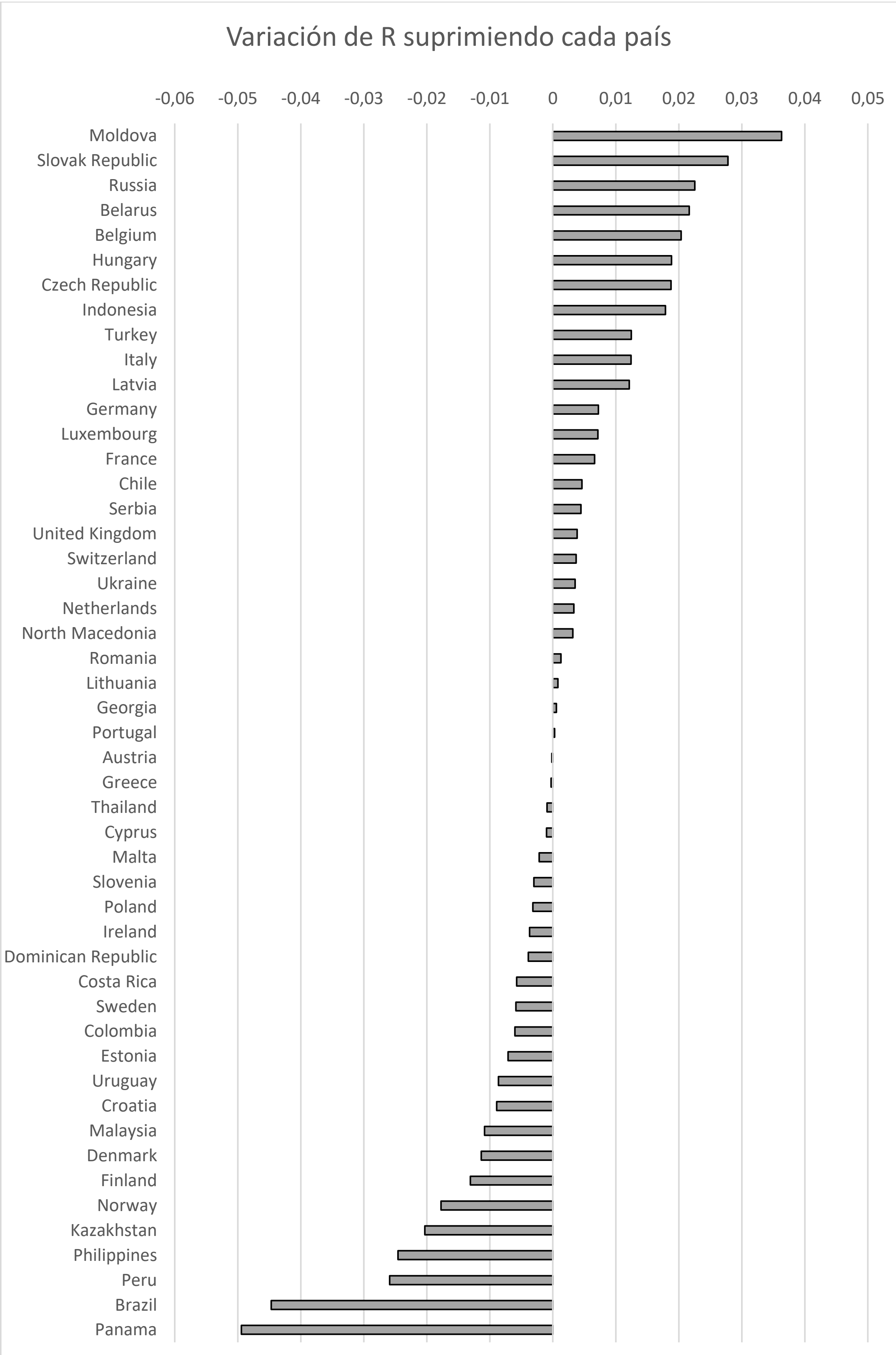
Tabla 6. Coeficientes de Gini para $\alpha = 10$ y variación de R cuando se suprime respectivamente cada Estado

Estado	Gini (BM)	Gini (PISA)	Variación R
Austria	30,8	34,8	-0,00014
Belgium	27,2	39,8	0,02038
Chile	44,4	36,3	0,00465
Colombia	50,4	39,3	-0,00602
Czech Republic	25,0	38,0	0,01877
Denmark	28,2	29,2	-0,01137
Estonia	30,3	24,4	-0,00711
Finland	27,3	29,2	-0,01306
France	32,4	40,1	0,00662
Germany	31,9	39,8	0,00723
Greece	32,9	34,8	-0,00028
Hungary	29,6	41,5	0,01886
Ireland	31,4	27,3	-0,00369
Italy	35,9	27,7	0,01241
Latvia	35,1	26,4	0,01215
Lithuania	35,7	34,6	0,00083
Luxembourg	35,4	46,4	0,00717
Netherlands*	28,1	35,6	0,00336
Norway	27,6	25,9	-0,01775
Poland	30,2	32,8	-0,00317
Portugal*	33,5	36,1	0,00030
Slovak Republic	25,0	39,8	0,02782
Slovenia	24,6	33,1	-0,00301
Sweden	30,0	30,8	-0,00586

Switzerland	33,1	39,3	0,00369
Turkey	41,9	33,5	0,01245
United Kingdom	35,1	31,0	0,00386
Belarus	25,2	38,7	0,02164
Brazil	53,9	44,2	-0,04467
Costa Rica	48,0	39,0	-0,00573
Croatia	29,7	28,1	-0,00890
Cyprus	32,7	30,4	-0,00101
Dominican Republic	43,7	38,2	-0,00387
Georgia	36,4	35,3	0,00056
Indonesia	37,8	28,5	0,01789
Kazakhstan	27,8	20,9	-0,02029
Malaysia	41,1	41,0	-0,01082
Malta	28,7	33,6	-0,00215
Moldova	25,7	41,8	0,03633
North Macedonia	33,0	38,8	0,00320
Panama	49,2	46,4	-0,04942
Peru	42,4	47,1	-0,02591
Philippines	42,3	46,5	-0,02457
Romania	35,8	43,5	0,00130
Russia	37,5	26,7	0,02254
Serbia	36,2	32,1	0,00445
Thailand	36,4	37,1	-0,00093
Ukraine	26,1	35,0	0,00355
Uruguay	39,7	41,1	-0,00864

Los resultados de la columna de la derecha de la tabla anterior se visualizan mejor en el gráfico siguiente. Puede verse, por ejemplo, que si se suprime Moldavia, la correlación sube 0,03633 puntos. Esto significa que, por así decir, este Estado *hace descender* la correlación en esa cantidad. Por el extremo contrario, si se suprime Panamá, la correlación desciende 0,04942 puntos, por lo que podemos entender que este país tira de la correlación hacia arriba en mayor medida. Esto se representa en el siguiente gráfico 6.

Gráfico 6. Variación en las Correlaciones de Pearson cuando se suprime cada Estado, de manera ordenada


Variación de R suprimiendo cada país
-0,06
-0,05
-0,04
-0,03
-0,02
-0,01
0
0,01
0,02
0,03
0,04
0,05
Moldova
Slovak Republic
Russia
Belarus
Belgium
Hungary
Czech Republic
Indonesia
Turkey
Italy
Latvia
Germany
Luxembourg
France
Chile
Serbia
United Kingdom
Switzerland
Ukraine
Netherlands
North Macedonia
Romania
Lithuania
Georgia
Portugal
Austria
Greece
Thailand
Cyprus
Malta
Slovenia
Poland
Ireland
Dominican Republic
Costa Rica
Sweden
Colombia
Estonia
Uruguay
Croatia
Malaysia
Denmark
Finland
Norway
Kazakhstan
Philippines
Peru
Brazil
Panama

La correlación de Pearson entre la columna de la desigualdad social y esta variación se eleva a -0,48970, lo que resulta muy significativo. El signo negativo se explica porque a mayor desigualdad, el Estado en cuestión hace crecer su correlación con los resultados de PISA, y por ello, al ser suprimida, la bajada de la correlación es mayor. De ahí el valor inverso. Se podría decir que no solo hemos demostrado que la desigualdad social correlaciona con la desigualdad educativa, sino que además la desigualdad social opera como un factor para que la desigualdad educativa sea mayor. Nos acercamos, pues, a la noción de causalidad.

Hay que advertir que en el gráfico anterior en los puestos superiores se concentran los antiguos países socialistas, que están generalmente ausentes en los inferiores, donde por otra parte hay mayor número de países americanos. Una hipótesis razonable sería que en aquellos países se generaron sistemas educativos más igualitarias que han pervivido incluso a la desaparición de los gobiernos de fidelidad soviética. De hecho, si suprimimos todos los antiguos países socialistas la correlación sube a R=0,5618. Obsérvese también que la elogiada Finlandia se encuentra entre los países que estimularían esta correlación entre desigualdad social y educativa.

Espero que ahora se entienda lo que dice al principio, que me proponía hacer una presentación científica, en sentido enfático, sobre la relación entre desigualdad social y desigualdad educativa. Naturalmente de esta presentación se podrían derivar muchas reflexiones. Entre otras, indagar las condiciones educativas no tanto de los países que obtienen mejores rendimientos, cuando de los que favorecen que descienda la correlación entre desigualdad social y educativa; también, profundizar en los sistemas educativos de los antiguos países socialistas, para ver sus diferencias específicas con otros sistemas, como los americanos. Intuyo que el desarrollo de la formación profesional o vocacional resulta un elemento distintivo, pero sería preciso una ratificación. Por último, una reflexión derivada es la necesidad de fundamentar los discursos sobre la educación menos en preceptos didácticos y más en análisis sociológicos, así como importar aquellas metodologías de investigación que han resultado contrastadas en otras disciplinas.

Bibliografía

Feynman (1974): Cargo Cult Science. Some remarks on science, pseudoscience, and learning how to not fool yourself. Disponible en: https://calteches.library.caltech.edu/51/2/CargoCult.htm (consulta 12/11/2021).

EUROSTAT (2021): Income quintile share ratio - S80/S20. Disponible en: https://ec.europa.eu/eurostat/databrowser/bookmark/f8b6b15e-e5d6-4c77-b114-7294a28444b7?lang=en (consulta 12/11/2021).

OECD (2019), PISA 2018 Results (Volume II): Where All Students Can Succeed, PISA, OECD Publishing, Paris, https://doi.org/10.1787/b5fd1b8f-en.

OECD (2021): Income inequality. Disponible en: https://data.oecd.org/inequality/income-inequality.htm (consulta 12/11/2021).

THE WORLD BANK (2021a), https://data.worldbank.org/indicator/SI.POV.GINI

THE WORLD BANK (2021b), http://iresearch.worldbank.org/PovcalNet/home.aspx (consulta 10/11/2021)