Análisis del Coeficiente de Gini: evidencia para el Estado de Puebla con un enfoque lineal

Analysis of the Gini Coefficient: evidence for the state of Puebla with a non-linear approach

*Emmanuel Olivera Pérez**

Resumen

El presente artículo tiene el objetivo de medir a través de Mínimos Cuadrados Ordinarios No Lineales (MCONL) el Coeficiente de Gini, siguiendo una diferente metodología para el diseño de la Curva de Lorenz para que ajuste mejor a los fines de la estimación. La estimación se basa en que la Curva de Lorenz puede simularse a través de una función de raíz n con coeficientes a y b de la forma $y = ax^b$, donde $0 \leq b \leq 1$ y el Coeficiente de Gini puede calcularse de la forma Gini=1-b. Los principales hallazgos usando datos de la encuesta EnBienComún del estado de Puebla 2015, son que la estimación de la desigualdad es mayor a la reportada por CONEVAL en 2015 para el estado de Puebla, dicha estimación es estadísticamente significativa. Los Criterios AIC y BIC resultan favorables para la especificación de la estimación en cuanto al tamaño de la muestra.

Palabras Clave: Coeficiente de Gini, Curva de Lorenz Invertida, Mínimos Cuadrados Ordinarios No Lineales, Intervalos Empíricos, Intervalos de Equidad.
Clasificación JEL: D63, D31, I32.

Abstract

The objective of this article is to measure the Gini Coefficient by Nonlinear Ordinary Least Squares (NLOLS), following a different methodology for the design of the Lorenz Curve so that it fits better for the purposes of estimation. The estimation is based on the fact that the Lorenz curve can be simulated through a function of root n with coefficients a and b of the form $y = ax^b$, where $0 \leq b \leq 1$ and the Gini Coefficient can be calculated in the form Gini = 1-b. The main findings using data from the EnBienComún survey of the State of Puebla 2015, are that the estimate of the inequality is greater than that reported by CONEVAL in 2015 for the

* Profesor Investigador. Universidad Popular Autónoma del Estado de Puebla (UPAEP), emmanuel.olivera@upaep.mx

State of Puebla, this estimate is statistically significant. AIC and BIC criteria are favorable for the specification of the estimate in terms of sample size.
Keywords: Gini Coefficient, Inverted Lorenz Curve, Nonlinear Ordinary Least Squares, Empirical Intervals, Equity Intervals.
JEL Code: D63, D31, I32.

1 Introducción

Las medidas de desigualdad se pueden dividir en dos grandes campos: las unidimensionales, que se basan en ingreso, salud, felicidad o utilidad y las multidimensionales que se basan en un análisis de múltiples atributos y factores (Adler, 2015).

Ambos enfoques han tenido grandes desarrollos, prácticamente en lo que respecta a las medidas unidimensionales, la propuesta de medición realizada por Gini en 1912, basada en el análisis de la Curva de Lorenz (Paglin, 1975) ha sido la más representativa y la más usada en este tipo de análisis y aunque en la literatura existen grandes aportaciones en cuanto a sus debilidades, esta medición sigue siendo parte importante en la medición de la desigualdad actualmente (Futing, 2006).

Pero ¿qué representan en realidad las medidas de desigualdad? Para Cowell y Feser (1996), las medidas de desigualdad representan un análisis numérico de las distribuciones empíricas del ingreso, en este sentido, este articulo tiene el objetivo de analizar la medida del coeficiente de Gini y estimarlo a través de una metodología diferente, la cual no está presente en la literatura actualmente, por lo que la principal contribución de este artículo es ofrecer una nueva metodología y evidencia de la medición del coeficiente de Gini con datos de la encuesta de EnBienComún del estado de Puebla 2015, encuesta de Movilidad Social.

Uno de los principales conflictos en la medida del coeficiente de Gini basados en la Curva de Lorenz es que la línea de perfecta equidad es un concepto teórico y no empírico, por lo que el remedio es estimar la línea de perfecta equidad por medio de una función propuesta a partir de la evidencia empírica.

Usando datos de ingresos familiares de los Estados Unidos de 1947 a 1964 se concluye que al estimar la línea de perfecta equidad mediante los ingresos anuales de las familias, se requiere la consideración de los ingresos por edad como aproximaciones de los ingresos de las trayectorias del ciclo de vida y que además es adecuado considerar los ingresos anuales como el principal factor de la estimación de la línea de equidad, con el objetivo de ajustar el coeficiente de Gini más realista y sin sesgo.

En 1976, Kakwani y Podder en su artículo *Efficient Estimation of the Lorenz Curve and Associated Inequality Measures from Grouped Observations* propone una estimación de la Curva de Lorenz mediante diferentes coordenadas a las tradicionales, con el objetivo de validar empíricamente que la desigualdad difiere al tratamiento de datos agrupados respecto a los no agrupados.

Se proponen cuatro métodos para estimar eficientemente la Curva de Lorenz: Mínimos Cuadrados Ordinarios, Mínimos Cuadrados Generalizados, Mínimos Cuadrados Directos y Mínimos Cuadrados Directos Asintóticos.

Con datos de Australia del estudio sobre Consumo, Finanzas y Gastos de 1967 a 1968 se concluye que al usar el test de la Bondad de Ajuste de los modelos estimados -la Bondad de Ajuste en este caso es una propuesta de Gastwirth y Smith (1972)-, el método de Mínimos Cuadrados Directos Asintóticos ofrece el mejor resultado (menores errores estándar).

En la publicación *Linear Measures of Income Inequality* (1976), Farhad Mehran propone que en general el coeficiente de Gini y la Desviación Media Relativa son medidas muy particulares de medidas lineales. Un resultado inicial es que solo el coeficiente de Gini cumple con el principio de transferencias Pigou-Dalton, pero la Desviación Media Relativa no, de tal forma que, a partir de este resultado, Mehran propone la siguiente familia de medidas lineales:

$$I = \frac{1}{\mu}\int_0^1 [F^{-1}(p) - \mu]W(p)dp \qquad [1]$$

De tal forma que para para cada p se defina una función de scores W(p) en la que resulta una medida lineal de desigualdad. El propósito es que estas medidas lineales impliquen fuertemente el principio de transferencias Pigou-Dalton, así que en términos de la Curva de Lorenz se puede expresar un nuevo índice lineal de desigualdad:

$$I = 6\int_0^1 [p - L(p) - \mu](1 - p)dp \qquad [2]$$

Paul Allison en 1978 publica *Measures of Inequality* donde realiza un análisis profundo de la validez de las medidas de desigualdad. Una de las líneas más importantes es analizar la valides de las comparaciones que se realizan entre países o ciudades, ya que bajo ciertos criterios dichas medidas no son comparables. Confirma que, si bien el coeficiente de Gini cumple con los criterios básicos de la Invariabilidad de Escala y el Principio de Transferencias, el Coeficiente de Variación y la Medida de Theil son preferibles.

Los resultados más importantes es que en general ninguna de estas medidas es comparable para datos a nivel de intervalos, a menos que se hagan sobre circunstancias especiales, en este sentido, es posible dicha comparación si se asume una distribución log-normal de los datos y usando una transformación logarítmica de la medida de desigualdad.

La introducción de una función de bienestar social como alternativa para derivar medidas de igualdad puede ser la solución a la búsqueda de medidas eficientes que ajusten a la distribución de los datos.

Particularmente desde finales de la década de los setenta, una serie de publicaciones han contribuido a mejorar las medidas de desigualdad, buscando un eficiente ajuste de los datos, tales como el uso de un enfoque de ingresos censurados a través de asumir una distribución truncada (Takayama, 1979), modificaciones eficientes al coeficiente de Gini como las medidas L-Gini, MP-Gini y P-Gini, esta última propuesta por Paglin (Formby y Terry, 1980). Un caso particular de análisis de la Curva de Lorenz en términos de considerar la elasticidad ingreso constante a lo largo de la distribución cuando se usan variables cualitativas, es validado por medio del cálculo de medidas de concentración, que se pueden medir a través de los gastos de los individuos (Krishnan, 1981). Scott Menar (1986) con el interés de validar la comparabilidad entre coeficientes de Gini, propone controlar por dimensiones como la extensión de la cobertura, tipo de población y tiempo, en este sentido las transformaciones lineales que se proponen de los coeficientes de Gini son correlaciones entre la distribución del ingreso y las dimensiones propuestas.

En 1988 se genera un aporte interesante. Satya Chakravarty confirma que el índice de Gini para medir desigualdad, se mantiene inalterado cuando se escalan los ingresos proporcionalmente, además de que resulta inalterado al multiplicar los ingresos por su media, que en tal caso tendríamos una medida de desigualdad absoluta.

Confirma también que en el caso de los ingresos individuales se puede considerar al índice de Gini como lineal, es decir, no existe una función de evaluación social estrictamente cóncava (Curva de Lorenz), en tal caso el coeficiente de Gini falla en satisfacer el axioma de

Transferencias Decrecientes, por lo que al realizar al ajuste del índice de Gini para satisfacer esta propiedad, resulta ser un índice normativo y que corresponde a una Función de Evaluación Social estrictamente cóncava.

En general la Curva de Lorenz asumiendo una distribución F queda:

$$L(p) = \frac{1}{\mu}\int_0^p [F^{-1}(t)]dt \qquad [3]$$

De este modo, asumiendo una función de Evaluación Social (W) que satisface el axioma de transferencias decrecientes tenemos:

$$W = \mu\left(2 - I_{\varphi}^{R}\right) \qquad [4]$$

Con $I_{\alpha}^{R} = 2[\int_0^1 \left(1 - L(p)\right)^{\alpha} dp]^{1/\alpha}$ [5]

En los noventas, las aportaciones más importantes están enfocadas a analizar las modificaciones de los coeficientes de Gini que validan sus principales resultados, por ejemplo, la valides de las propiedades asintóticas de las medidas relativas y absolutas del coeficiente de Gini (medidas generalizadas del coeficiente de Gini) en términos de la Curva de Lorenz ordenada (Barrett y Pendakur, 1995).

En este mismo sentido, con el objetivo de dar robustez a las medidas de desigualdad, en particular la del coeficiente de Gini, Cowell y Feser (1996) realizan tanto estimaciones paramétricas como no paramétricas a través de simulaciones Pareto, Log-Normales, Gamma entre otras. En 1999 un interesante artículo cuyo objetivo fue reunir 21 medidas de desigualdad (incluida el coeficiente de Gini) para analizar los criterios de desigualdad que se persiguen y el posible sesgo que conlleva cada una de ellas.

El resultado general es que los criterios de análisis (asumiendo una familia de distribuciones no normales) dan respuesta a que las diferencias existentes tienen que ver con el tipo de desigualdad que se pretende analizar y con el sesgo que la distribución particular de los ingresos que conlleve.

Extensiones interesantes que analizan versiones diferentes del coeficiente de Gini, por un lado, considerando la estratificación social -clases- (Futing, 2006) y por el otro al considerar de forma particular la desigualdad existente entre los ingresos más altos y el coeficiente de Gini, de tal forma que el coeficiente de Gini tiende a ser más significativo (Leigh, 2007).

En 2008, John Golden en su artículo *A Simple Geometric Approach to Approximating the Gini Coeffcient* presenta una revisión y una interesante medición empírica del coeficiente de Gini a través del cálculo empírico de la Curva de Lorenz (*Z-Gradient Rule*).

La estimación empírica consiste en una transformación por quintiles de 621 observaciones de Ingresos. El artículo destaca que es usual que el coeficiente de Gini en muchas ocasiones es mejorado por la regla del Trapezoide.

Otro artículo interesante, muy relacionado al de Golden (2008) analiza dos formas importantes de medir el coeficiente de Gini, la primera de ellas ajustando los percentiles de la distribución en relación al promedio de ingresos, y el segundo en términos de la selección aleatoria de los ingresos más bajos respecto a la media del ingreso.

Para ello propone una familia de funciones de Pareto que aproximan a la Curva de Lorenz mediante la función de Poder:

$$L(p) = ap + (1 - a)p^b \qquad [6]$$

De forma más reciente, dos publicaciones tienen en el centro de su análisis la estimación empírica de la Curva de Lorenz, la primera de ellas (Jantzen, 2012) determina un modelo de dos parámetros para la Curva de Lorenz, basado en la interpolación entre las similitudes en la dinámica de concentración entre los más bajos y más altos ingresos, generando dos índices de desigualdad que describen las posibles coincidencias o las posibles diferencias.

Dichos índices en términos de Jantzen son suficientes para describir la desigualdad en la distribución. Usando datos de ingreso del 2009 para Estados Unidos se propone que a partir del coeficiente de Gini en términos de la Curva de Lorenz:

$$G = 2\int_0^1 x - L(x)dx \qquad [7]$$

Se realiza una estimación polinómica de grado 5 (5 coordenadas=quintiles):

$$L(x) = a_1x + a_2x^2 + a_3x^3 + a_4x^4 + a_5x^5 \qquad [8]$$

Asumiendo la similitud de las esquinas extremas de la distribución de la forma:

$$L(x) = 1 - (1 - x)^q \qquad [9]$$

$$L(x) = cx^q \qquad [10]$$

Tal que se construyó un modelo hibrido de la forma:

$$L(x) = x^p(1 - (1 - x)^q) \qquad [11]$$

De lo que se concluye que las principales ventajas son: ajuste excelente a lo largo de toda la distribución, distinción entre la intersección de la Curva de Lorenz con idénticos índices de Gini, se requieren dos parámetros y mejora las interpretaciones del coeficiente de Gini.

La segunda publicación de interés es la realizada por Tillé y Langel (2012), plantea que la medición de la Curva de Lorenz esta sobreestimada cuando se considera la variabilidad al interior de las clases, por lo que el coeficiente de Gini resulta subestimado.

La propuesta para reducir este sesgo, es considerar un estimador estrictamente convexo basado en una interpolación por aproximaciones de la función de distribución de quintiles acumulada, de tal forma que integrando la función de aproximación de la Curva de Lorenz puede obtenerse el coeficiente de Gini considerando ahora la variabilidad intra-clase.

En este caso platea que a través de la función que da origen a la Curva de Lorenz:

$$\int_0^1 Q(p)dp = \int_0^\infty xdF(x) = \mu \qquad [12]$$

Donde F es la función cuantílica y derivando se obtiene la trayectoria de la Curva de Lorenz:

$$\frac{dL(\alpha)}{d\alpha} = \frac{Q(\alpha)}{\mu} \qquad [13]$$

Finalmente se estima el coeficiente de Gini de la forma:

$$G = 1 - 2\int_0^1 L(\alpha)d\alpha \qquad [14]$$

2 Métodos

Asumimos que existe una distribución de ingreso x, de tal forma que su función de densidad es:

$$F(x) = \int_0^\infty f(x)\,dx \qquad [15]$$

Para efectos del cálculo del coeficiente de Gini, la Curva de Lorenz se conforma por coordenadas de la distribución acumulada del ingreso y la distribución acumulada de la población, otra alternativa es tomar cuantiles de la distribución.

Para los fines del artículo, nos interesa generar dos medidas particulares: una por n cuantiles que representen la distribución del ingreso y otra por n cuantiles que representen la distribución equitativa del ingreso.

La primera medida está relacionada con la distribución de los ingresos, es decir, tomamos n cuantiles empíricamente establecidos de la forma:

$$q_j = \frac{j}{100}(n+1) \qquad [16]$$

Con j=1, 2,……, k cuantiles, tal que para cada q_j se obtengan los siguientes k intervalos para una distribución ordenada x:

$$[x_{q_{j-1}}, x_{q_j}] \qquad [17]$$

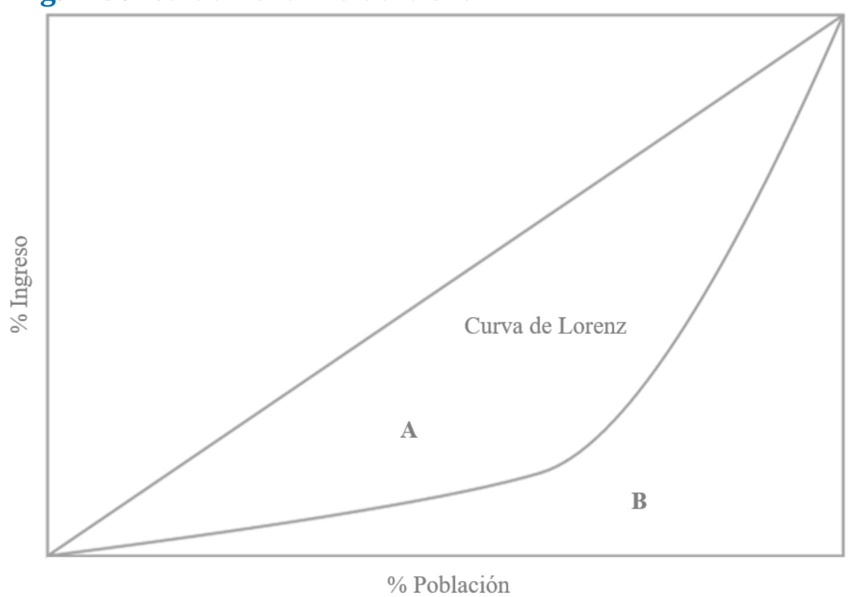
De igual forma, definimos ahora los cuantiles de equidad, que son teóricamente diseñados, partiendo de que conocemos el rango de la distribución R, de tal forma que para un número determinados de cuantiles k, tenemos los siguientes intervalos:

$$\left[x_{min}, x_{min} + \frac{R}{k}\right], \left(x_{min} + \frac{R}{k}, x_{min} + 2\frac{R}{k}\right], \left(x_{min} + 2\frac{R}{k}, x_{min} + 3\frac{R}{k}\right], \dots \dots, \qquad [18]$$

$$\left(x_{min} + (k-1)\frac{R}{k}, x_{min} + k\frac{R}{k}\right] \qquad [19]$$

Así, combinando los dos cuantiles formamos una caja cuantílica por intervalos en x y no en porcentajes como tradicionalmente se mide el coeficiente de Gini (ver a continuación):

Fig. 1 Curva de Lorenz tradicional



Fuente: elaboración propia.

De aquí que simplemente el coeficiente de Gini de forma aritmética resulta en medir la desigualdad de la siguiente forma:

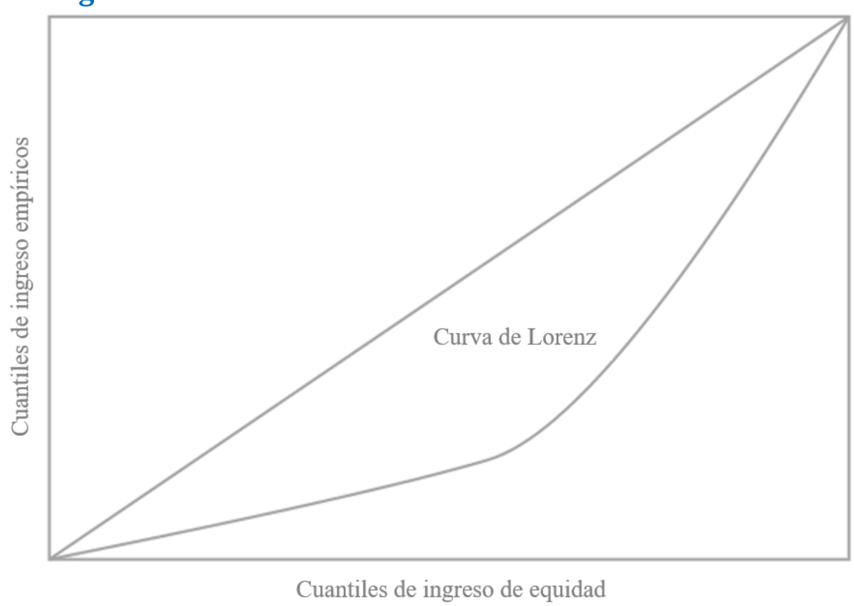
$$Gini = \frac{A}{A+B} \qquad [20]$$

Donde sí Gini==1 implica perfecta igualdad, dado que el resultado es A/A, y si Gini==0 implica que A=0, que implica perfecta desigualdad.

La propuesta entonces es tomar los intervalos en vez de los porcentajes, aun cuando el resultado de la Curva de Lorenz resulta similar, la diferencia significativa es que no medimos el porcentaje de la población que se encuentra en cada intervalo de ingreso de la distribución, sino medimos como referencia, los intervalos en los que la distribución de ingreso es equitativa, esto también nos permite dibujar la recta de perfecta equidad en la distribución del ingreso.

De tal forma que resulta una caja cuyos extremos son las coordenadas (x_{min}, x_{min}) y (x_{max}, x_{max}), y en ambos ejes son los intervalos de ingreso medidos en diferentes posiciones de la distribución, por un lado, la empírica y por otro lado una teórica, como lo demuestra la gráfica 2:

Fig. 2. Curva de Lorenz basada en cuantiles



Fuente: elaboración propia.

La Función de Desigualdad no es más que la función de la Curva de Lorenz, pero invertida, es decir, simplemente si invertimos los ejes, la Curva de Lorenz invertida es la curva de desigualdad que nos interesa analizar.

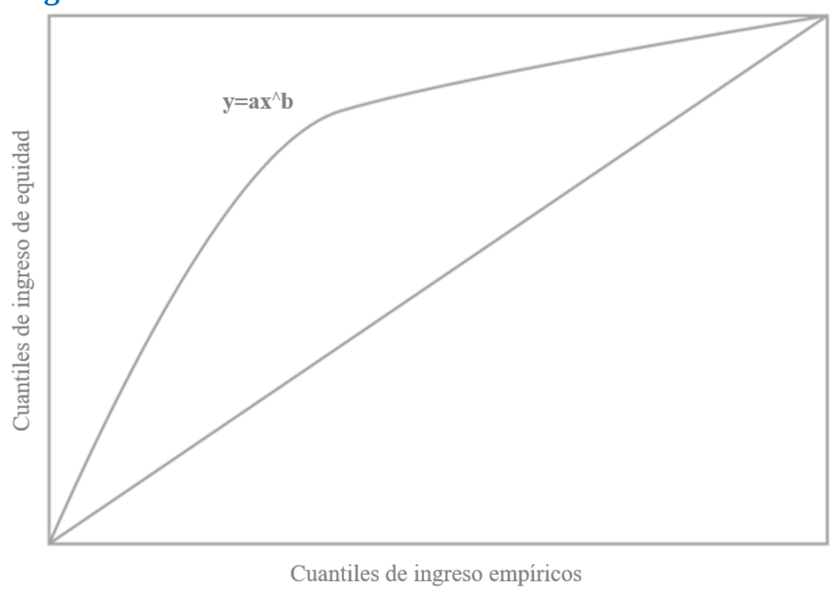
Dada la forma de la función (Curva de Lorenz invertida), podemos aproximar su forma a la familia de funciones de Potencias con exponentes racionales positivos:

$$y = ax^b \qquad [22]$$

Donde a ≥ 0 y 0 ≤ b ≤ 1

De forma específica es la familia de funciones de Raíz Positiva con coeficiente positivo. Dado que la Curva de Lorenz parte del origen, la función que imita su dinámica no tiene intercepto. De tal forma que tenemos la siguiente definición de la Curva de Lorenz invertida.

Fig. 3 Curva de Lorenz invertida



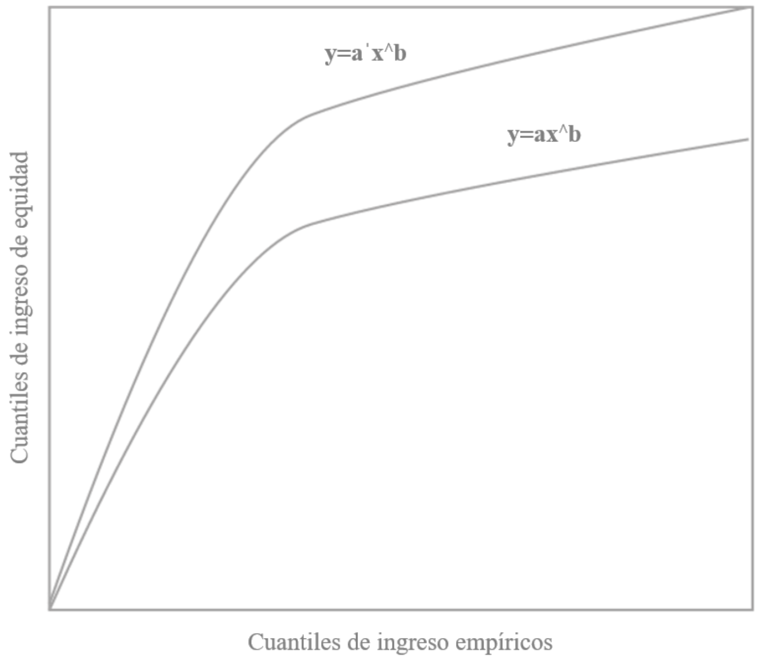
Fuente: elaboración propia.

Con esta definición de la Curva de Lorenz en fácil ver que la función f(x)=y tiene ciertos atributos que ajustan idealmente a la función de la Curva de Lorenz.

$$f(x) = y \begin{cases} Dominio = [x_{q_{j-1}}, x_{q_j}] \\ Codominio = [x_{qe_{j-1}}, x_{qe_j}] \end{cases} \qquad [23]$$

Donde $[x_{q_{j-1}}, x_{q_j}]$ son los intervalos empíricos de la distribución de ingreso, mientras que los intervalos $[x_{qe_{j-1}}, x_{qe_j}]$ son llamados intervalos de igualdad. De forma gráfica podemos ver que las implicaciones del parámetro a, cuando a’>a:

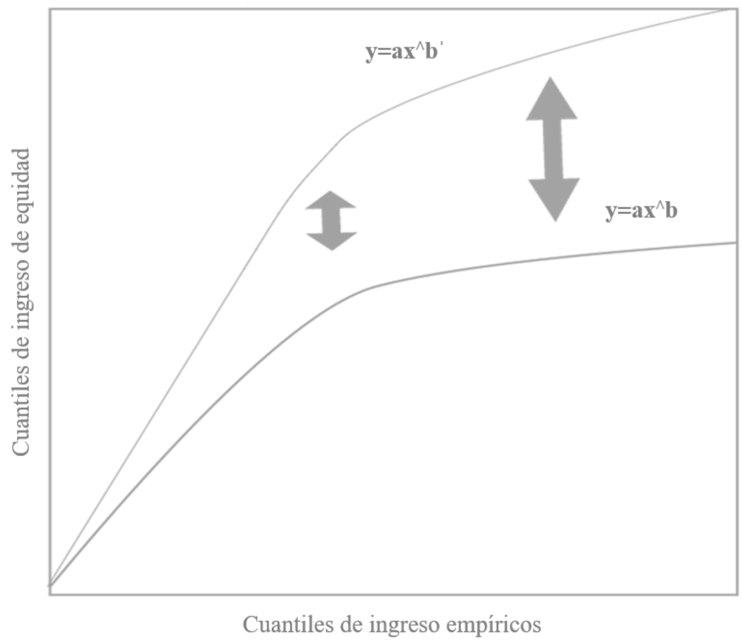
Fig. 4 Diferentes pendientes de la Curva de Lorenz



Fuente: elaboración propia.

El parámetro b es aún más fácil de modelar, ya que mide la pendiente de la Curva de Lorenz invertida, si b’>b:

Fig. 5 Diferentes exponentes de la Curva de Lorenz



Fuente: elaboración propia.

De tal forma que mientras el parámetro a desplaza a la función de forma proporcional a cada valor de x, el parámetro b tiende a cambiar la pendiente de la función, incrementando la distancia entre las funciones con b’ y b.

De forma particular el parámetro b mide el grado de inclinación de la trayectoria de la Curva de Lorenz invertida, mientras que el parámetro a mide el grado en el que se desplaza la curva, de tal forma que a continuación se resumen los efectos combinados de a y b para efectos de dos funciones de desigualdad:

$$y = a + x^b \qquad [24]$$

$$y' = a' + x^{b'} \qquad [25]$$

Dadas las diferentes relaciones entre los parámetros podemos resumir lo siguiente:

∃ a, a’, b, b’ tal que:

$$Desigualdad = \begin{cases} a' > a, \ and b \rightarrow 0 \\ a' < a, \ and b' \rightarrow 0 \end{cases} \qquad [26]$$

De tal forma que, dadas las relaciones entre los coeficientes, el parámetro b define el nivel de desigualdad existente en la distribución de ingreso, en este caso, para definir igualdad, se tiene que las siguientes relaciones son válidas:

$$Igualdad = \begin{cases} a' > a, \ and b' \rightarrow 1 \\ a' < a, \ and b \rightarrow 1 \end{cases} \qquad [27]$$

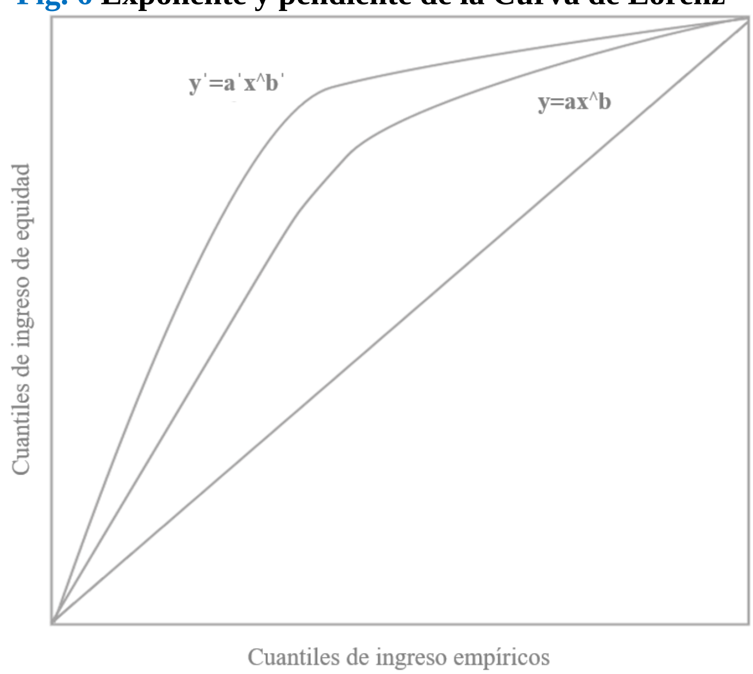
Así que hemos llegado a la siguiente definición: el parámetro b es una medida de la desigualdad en la distribución del ingreso, medida que es acotada entro 0 y 1 donde:

$1 \rightarrow Maxima\ Igualdad$

$0 \rightarrow Maxima\ Desigualdad$

Gráficamente se observa las siguientes relaciones entre las curvas de desigualdad y, y’:

Fig. 6 Exponente y pendiente de la Curva de Lorenz



Fuente: elaboración propia.

De tal forma que el coeficiente de Gini queda definido como:

$$Gini = 1 - b \qquad [28]$$

Donde ahora tenemos la misma definición que tradicionalmente tenemos del coeficiente de Gini:

$0 \rightarrow Maxima\ Igualdad$

$1 \rightarrow Maxima\ Desigualdad$

La propuesta de estimación de la Curva de Lorenz es una técnica poco usada en la literatura económica convencional, no por la complejidad que implica, ni por resultar una técnica trivial, sino más bien porque las aplicaciones son en sentido estricto muy pequeñas, ya que, en términos más menos generales, siempre es posible linealizar una función y proceder con el ya tradicional Mínimos Cuadrados Ordinarios.

Sin embargo, linealizar las funciones casi siempre nos lleva a utilizar las propiedades de los logaritmos, que tienden a cambiar las unidades de las variables y entonces como es de esperarse, eso no implica una transformación monótona que contribuya a mantener las propiedades iniciales de las funciones a explicar. MCONL se basa en asumir la siguiente función:

$$y_i = g(X_i, B) + e_i \qquad [29]$$

Donde g es una función no lineal que depende del vector de variables explicativas y del vector de coeficientes a estimar, y de forma idéntica consideramos el término de error, el cual usualmente tiende a ser lineal, aunque no necesariamente. De acuerdo a la definición anterior, el modelo lineal es solo un caso especial de $g(X, B)$ (Greene, 2008). Los supuestos en su forma reducida de MCONL son: 1) Identificación de los parámetros; 2) Media Condicional de Cero; 3) Homoscedasticidad; 4) No Autocorrelación Serial

Con tales supuestos llegamos a estimadores insesgados y de mínima varianza, es decir, son consistentes y eficientes.

Así que, de la misma forma en la que definimos la función de cuadrados en MCO, definimos la función de Cuadrados de la siguiente forma en la versión No Lineal:

$$e^2 = E(y_i - g(X_i, B))^2 \qquad [30]$$

Las condiciones de primer orden no tienen una solución explicita, por lo que en general el modelo de Mínimos Cuadrados No Lineales está definido no en el sentido No lineal de la función $g(X, B)$, sino más bien en el sentido de lo no linealidad de las funciones obtenidas como resultado de las condiciones de primer orden.

Dado lo anterior, la función $g(X, B)$ puede simplificarse por una aproximación de Taylor, calculando la derivada parcial k del parámetro m x_k^m:

$$\frac{\partial g(X,B)}{\partial b_k^m} \qquad [31]$$

De tal forma que la definición final para conocer al vector de parámetros B puede expresarse con base a b^m, es decir el parámetro m del vector B:

$$g(X, B) \approx [g^m - \sum_{k=1}^{K} x_k^m b_k^m] + \sum_{k=1}^{K} x_k^m B_k \qquad [32]$$

Donde ahora podemos definir el término de error de la siguiente forma:

$$e^m = e + [g(X, B) - \{g^m - \sum_{k=1}^{K} x_k^m b_k^m + \sum_{k=1}^{K} x_k^m B_k\}] \qquad [33]$$

3 Resultados

La base de datos utilizada en este artículo, es una encuesta sobre Movilidad Social realizada en el 2015 llamada EnBienComún para el estado de Puebla, tiene un total de 704 variables y un total de observaciones de 3,003. Para nuestra aplicación se han elegido 18 variables que consiste en evaluar 17 activos (variables binarias) del hogar del entrevistado y el ingreso mensual del individuo (pesos).

En esta sección estimamos la Curva de Lorenz, en específico el coeficiente b, de tal forma que podamos medir el coeficiente de Gini de acuerdo a la definición realizada en la sección anterior.

Según el Consejo Nacional de Evaluación de la Política Social (CONEVAL), la cual es una institución oficial cuyo objetivo es la medición de la pobreza y desigualdad en México, determino que el coeficiente de Gini para el estado de Puebla era de 0.482 en 2010, cifra que aumentó en 2012 con 0.485, en 2014 se ubicó en 0.572, la última cifra para el 2016 fue de 0.439. México en este contexto tuvo un coeficiente de Gini para los mismos años de 0.505, 0.509, 0.498, 0.503 y 0.498 respectivamente.

A continuación, se muestra los estadísticos básicos del ingreso mensual en Puebla:

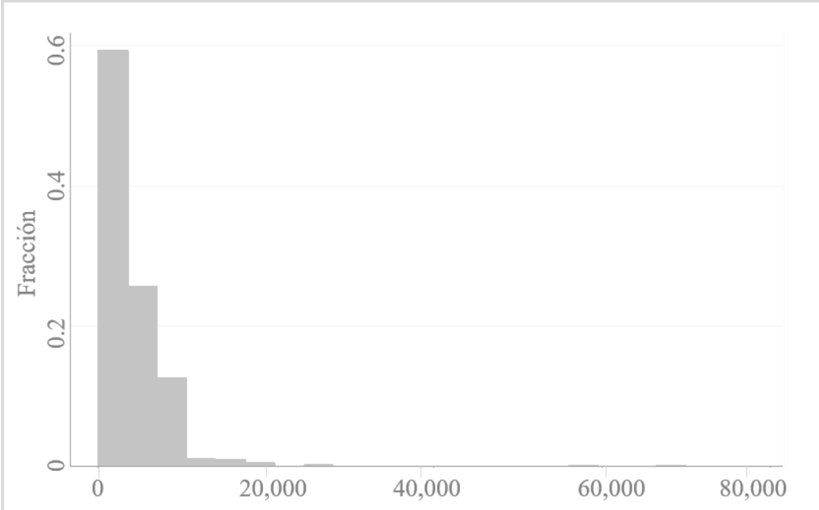
Tabla 1 Estadísticos básicos del ingreso mensual (pesos) de Puebla 2015

Variable	Obs	Mean	Std. Dev.	Min	Max
Ingreso mensual	1,492	3847.7	3934.46	0	70000

Fuente: elaboración propia con datos de la encuesta EnBienComun Puebla 2015.

Se puede visualizar que, de 3003 observaciones, solo se cuenta con 1492 observaciones en cuanto a ingreso mensual. A continuación, se muestra la distribución del ingreso, la cual está claramente sesgada a la derecha:

Graf. 7 Distribución de ingreso para el estado de Puebla 2015



Fuente: elaboración propia.

Tomando la definición de la función de desigualdad:

$$y = ax^b \qquad [34]$$

Donde:

$$\begin{cases} y = itervalos\ de\ equidad\ del\ ingreso \\ x = intervalos\ empiricos\ del\ ingreso \end{cases}$$

El modelo en este sentido puede expresarse de la siguiente forma para efectos de la estimación:

$$y_i = ax_i^b + e_i \qquad [35]$$

Partiendo la distribución del ingreso en percentiles para generar los intervalos empíricos, obtenemos la siguiente distribución:

Tabla 2 Distribución de percentiles del ingreso

Percentil	Frecuencia	%	Frec. Acum.	Percentil	Frecuencia	%	Frec. Acum.
1	18	1.21	1.21	58	26	1.74	58.78
2	22	1.47	2.68	59	4	0.27	59.05
3	12	0.8	3.49	60	41	2.75	61.8
4	17	1.14	4.62	62	21	1.41	63.2
5	11	0.74	5.36	64	98	6.57	69.77
6	33	2.21	7.57	70	8	0.54	70.31

8	12	0.8	8.38	71	28	1.88	72.18
9	107	7.17	15.55	73	31	2.08	74.26
16	56	3.75	19.3	75	65	4.36	78.62
20	51	3.42	22.72	79	8	0.54	79.16
23	11	0.74	23.46	80	76	5.09	84.25
24	73	4.89	28.35	85	25	1.68	85.92
29	21	1.41	29.76	86	2	0.13	86.06
30	26	1.74	31.5	87	34	2.28	88.34
32	129	8.65	40.15	89	12	0.8	89.14
41	15	1.01	41.15	90	30	2.01	91.15
42	31	2.08	43.23	92	72	4.83	95.98
44	58	3.89	47.12	96	22	1.47	97.45
48	42	2.82	49.93	98	10	0.67	98.12
50	3	0.2	50.13	99	15	1.01	99.13
51	103	6.9	57.04	100	13	0.87	100
				Total	1492	100.00	

Fuente: elaboración propia con datos de la encuesta EnBienComun Puebla 2015.

De tal forma que los Intervalos de Equidad y los Intervalos Empíricos quedan de la siguiente forma:

Tabla 3 Intervalos de ingreso mensual en pesos del estado de Puebla por percentiles

Percentil	Intervalos empíricos	Intervalos de equidad	Percentil	Intervalos empíricos	Intervalos de equidad
/	0	0	58	3200	36666.6667
1	0	1666.66667	59	3300	38333.3333
2	300	3333.33333	60	3500	40000
3	400	5000	62	3600	41666.6667
4	500	6666.66667	64	4000	43333.3333
5	600	8333.33333	70	4200	45000
6	800	10000	71	4500	46666.6667
8	900	11666.6667	73	4800	48333.3333
9	999	13333.3333	75	5000	50000
16	1000	15000	79	5500	51666.6667
20	1200	16666.6667	80	6000	53333.3333

Cuadro 3. Intervalos de ingreso mensual en pesos del estado de Puebla por percentiles (continuación)

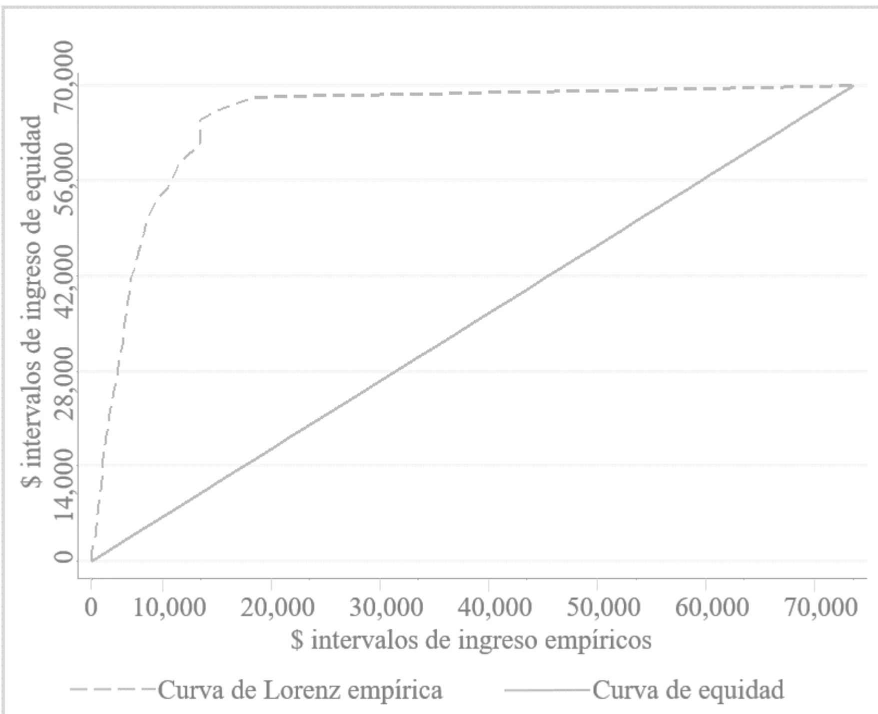
Percentil	Intervalos empíricos	Intervalos de equidad	Percentil	Intervalos empíricos	Intervalos de equidad
23	1300	18333.3333	85	7000	55000
24	1500	20000	86	7500	56666.6667
29	1600	21666.6667	87	8000	58333.3333
30	1800	23333.3333	89	9000	60000

32	2000	25000	90	9998	61666.6667
41	2300	26666.6667	92	9999	63333.3333
42	2400	28333.3333	96	10000	65000
44	2500	30000	98	12000	66666.6667
48	2800	31666.6667	99	15000	68333.3333
50	2940	33333.3333	100	70000	70000
51	3000	35000			

Fuente: elaboración propia con datos de la encuesta EnBienComun Puebla 2015.

La siguiente gráfica muestra las diferentes relaciones entre los intervalos, donde se incluyó la llamada línea de equidad.

Fig. 8 Curva de Lorenz invertida para el estado de Puebla 2015



Fuente: elaboración propia.

Como puede verse en la gráfica, tenemos la Curva de Lorenz invertida y la línea de perfecta equidad en el cuadrante con limites (0, 0), (0, 70000), (70000, 0) y (70000, 70000).

Los resultados del modelo son los siguientes:

Tabla 4 **Resultados del modelo**

Parámetros estimados (MCONL)	
a	2,321.29
	[630.399]**
b	0.339
	[0.030]**
R2	0.95
N	43
*** p<0.05; ** p<0.01**	

Nota. Resultados basados en una estimación por MCONO usando la encuesta EnBienComun Puebla 2015.
Fuente: elaboración propia.

Podemos observar que la estimación corresponde a una elevada desigualdad en la distribución del ingreso (ver Curva de Lorenz invertida) de tal forma que el parámetro a tiende a ser muy alto y en consecuencia b tiende a ser muy pequeño (ver relaciones en el apartado 2).

La R^2 tiende a ser muy elevada, lo que es un resultado esperado ya que la función propuesta ajusta a la dinámica de la Curva de Lorenz. El número de observaciones (intervalos) se ajustó a 43 debido a la ausencia de observaciones en los intervalos de la distribución del ingreso además de que en la distribución hay valores repetidos en los datos de ingreso en la muestra.

El Coeficiente de Gini se calcula de la siguiente forma:

$$Gini = 1 - b = 1 - 0.339 = 0.661 \qquad [36]$$

Podemos apreciar que el coeficiente de Gini está muy por encima de lo que reporta CONEVAL para el estado de Puebla (0.572 en 2014 y 0.439 en 2016), de tal forma que, según la evidencia empírica, tomando datos del 2015, la propuesta de medición de este articulo refleja una mayor desigualdad hasta de 0.1555 puntos respecto al promedio del coeficiente de Gini entre 2014 y 2016 (el promedio es de 0.5055).

En sentido estricto las estimaciones entre CONEVAL y la del presente articulo deben de coincidir dado que son muestras representativas del estado de Puebla, sin embargo, dadas las diferentes metodologías podemos justificar la diferencia de resultados.

La siguiente tabla muestra diferentes resultados del Criterio AIC (*Akaike Information Criteria*) y del Criterio BIC (*Bayesian Information Criteria*) con base a diferentes muestras. El interés es validar el modelo propuesto en términos del número de observaciones (intervalos de ingreso considerados en la distribución), lo que contribuye a determinar que no hay ganancia significativa al considerar un mayor número de observaciones, ya que, aunque AIC no disminuye (sensible al número de variables), el BIC aumenta, lo que implica un menor ajuste empírico del modelo.

Tabla 5 Criterios de Información de Akaike, Schwarz (AIC) y Bayesiano (BIC) para diferentes tamaños muéstrales

Modelo	Obs	ll(modelo)	g.l.	AIC	BIC
1	10	-454.1846	2	912.3692	912.9744
2	43	-454.1846	2	912.3692	915.8916
3	100	-454.1846	2	912.3692	917.5796
4	1,000	-454.1846	2	912.3692	922.1847
5	10,000	-454.1846	2	912.3692	926.7899
6	100,000	-454.1846	2	912.3692	931.3951

Nota. Resultados de varios modelos basados en una estimación por MCONO usando la encuesta EnBienComun Puebla 2015.
Fuente: elaboración propia.

4 Conclusiones

Este artículo ha considerado la Curva de Lorenz, construida a partir de intervalos de ingreso empíricos y teóricos (eficientes) de la distribución de ingreso para el estado de Puebla. La estimación se ha realizado por Mínimos Cuadrados Ordinarios No Lineales (MCONL), asumiendo una función no lineal que ajusta a la forma de la Curva de Lorenz con parámetros a, b. En términos de los resultados obtenidos se concluye lo siguiente:

La construcción de la Curva de Lorenz invertida a partir de los llamados intervalos empíricos y eficientes no altera los supuestos o resultados de la Curva de Lorenz, ya que se preservan todas las propiedades teóricas y técnicas de la forma de la curva.

La propuesta de la función $Y=aX^b$ para estimar la Curva de Lorenz invertida resulta ser una aproximación ideal en términos de los resultados obtenidos respecto a la trayectoria de desigualdad de la distribución de ingreso.

Mínimos Cuadrados Ordinarios No Lineales es un método que ajusta empíricamente a la función propuesta con el objetivo de estimar los parámetros de interés. En este sentido la medición del coeficiente de Gini es coherente con los métodos de medición previamente revisados en la literatura, por lo que resulta en una estimación del coeficiente de Gini comparable con el resto de las mediciones empíricas.

El cálculo del coeficiente de Gini en este artículo apunta a una mayor desigualdad en el Estado de Puebla de lo que las cifras oficiales han declarado para el 2015. Esta medición del coeficiente de Gini es coherente con otros análisis, además de consistentes y eficientes, lo que permite compararlos con otras mediciones. Desde este punto de vista la importancia empírica de los parámetros estimados a partir de la función de ajuste propuesta para la Curva de Lorenz es crucial para el análisis.

El estudio presenta ciertas limitaciones en términos del número de observaciones consideradas, sin embargo, proponiendo diferentes tamaños de la muestra, se observa que no hay ganancias en el modelo propuesto incluyendo un mayor número de observaciones, al menos por el resultado del Criterio de Bayes.

Bibliografía

Aaron, S., y McGuireciety, M. (1970). Public Goods and Income Distribution. *Econometrica*, *38*(6), 907–920.

Allison, P. D. (2018). Inequality Measures for Nominal Data. *American Sociological Review*, *46*(3), 371–373.

Allison, P. D. (1978). Measures of Inequality. *American Sociological Review*, *43*(6), 865–880.

Armour, B. P., Burkhauser, R. V, y Larrimore, J. (2013). Deconstructing Income and Income Inequality Measures : A Crosswalk from Market Income to Comprehensive Income. *The American Economic Review*, *103*(3), 173–177.

Barrett, G., y Pendakur, K. (1995). The Asymptotic Distribution of the Generalized Gini Indices of Inequality. *The Canadian Journal of Economics*, *28*(4b), 1042–1055.

Braun, D. (1988). Multiple Measurements of U. S. Income Inequality. *The Review of Economics and Statistics*, *70*(3), 398–405.

Buchanan, A. (1986). Inequality. *Philosophy and Public Affairs*, *15*(2), 99–121. https://doi.org/10.3982/ECTA6248

Chakravarty, S. R. (1988). Extended Gini Indices of Inequality. *International Economic Review*, *29*(1), 147–156.

Champernowne, D. G. (1974). A Comparison of Measures of Inequality of Income Distribution. *The Economic Journal*, *84*(336), 3787–3816.

CONEVAL. (2009). *Metodología para la medición multidimensional de la pobreza en México*. (CONEVAL, Ed.), *CONEVAL*, *México DF*. (Primera Ed). México: CONEVAL.

Corrado, G. (1921). Measurement of Inequality of Incomes. *The Economic Journal*, *31*(121), 124–126. https://doi.org/10.1111/j.l468-0297.2012.02500.x.

Cowell, F. A., y Jenkins, S. P. (1995). How Much Inequality Can we Explain ? A Methodology and an Application to the United States. *The Economic Journal*, *105*(429), 421–430.

Cowell, F. A., y Victoria-Feser, M.P. (1996). Robustness Properties of Inequality Measures. *Econometrica*, *64*(1), 77–101.

Dalton, H. (1920). The Measurement of the Inequality of Incomes. *The Economic Journal*, *30*(119), 348–361. https://doi.org/10.1016/j.surg.2014.04.001

Ebert, U. (1999). Dual Decomposable Inequality Measures. *The Canadian Journal of Economics*, *32*(1), 234–246.

Éltetö, S. Ö., y Frigyesociety, E. (1968). New Income Inequality Measures as Efficient Tools for Causal Analysis and Planning. *Econometrica*, *36*(2), 383–396.

Farnsworth, D., y Orr, R. (1986). Gini Means. *The American Mathematical Monthly*, *93*(8), 603–607.

Farris, F. A. (2010). The GINI index and measures of inequality. *American Mathematical Monthly*, *117*(10), 851–864. https://doi.org/10.4169/000298910X523344

Fields, G. S., y Efe, A. (1999). Measuring Movement of Incomes. *Economica*, *66*(264), 455–471.

Frees, E., Meyers, G., y Cummings, D. (2011). Summarizing Insurance Scores Using a Gini Index. *Journal of the American Statistical Association*, *106*(495), 1085–1098. https://doi.org/10.1198/jasa.2011

Futing, T. (2006). Measuring and Analyzing Class Inequality with the Gini Index Informed by Model-Based Clustering. *Sociological Methodology*, *36*(2006), 201–224.

Garner, T. I. (1993). Consumer Expenditures and Inequality : An Analysis Based on Decomposition of the Gini Coefficient. *The Review of Economics and Statistics*, *75*(1), 134–138.

Gary S. Fields. (1979). Decomposing LDC Inequality. *Oxford Economic Papers*, *31*(3), 437–459.

Gastwirth, J. L. (1974). Large Sample Theory of Some Measures of Income Inequality. *Econometrica*, *42*(1), 191–196.

Gini, C. (1912). Variabilità e Mutabilità (Variability and Mutability). *Memorie Di Metodologica Statistica (Ed. Pizetti E, Salvemini, T)*, 156. https://doi.org/10.1037/h0079220

Golden, J. (2008). A Simple Geometric Approach to Approximating the Gini Coefficient. *The Journal of Economic Education*, *39*(1), 68–77.

Gruen, C., y Klasen, S. (2008). Growth, inequality, and welfare: Comparisons across space and time. *Oxford Economic Papers*, *60*(2), 212–236. https://doi.org/10.1093/oep/gpm042

Guillermina, J. (1980). A New Theory of Distributive Justice. *American Sociological Review*, *45*(1), 3–32.

Haughton, J., y Khandker, S. (2009). *Handbook on Poverty and Inequality*. (T. W. B. Group, Ed.) (First Edition). Washington, DC: The World Bank. https://doi.org/10.1596/978-0-8213-7613-3

Hodler, R. (2009). Redistribution and Inequality in a Heterogeneous Society. *Economica*, *76*(304), 704–718.

Jacobs, D. (1980). Dimensions of Inequality and Public Policy in the States. *The Journal of Politics*, *42*(1), 291–306.

Jantzen, R. T., y Volpert, K. (2012). On the mathematics of income inequality: Splitting the gini index in two. *American Mathematical Monthly*, *119*(10), 824–837. https://doi.org/10.4169/amer.math.monthly.119.10.824

Jenkins, S. P., Burkhauser, R. V., Feng, S., y Larrimore, J. (2011). Measuring inequality using censored data : a multiple-imputation approach to estimation and inference. *Journal of the Statistical Society*, *174*(1), 63–81.

Johannesson, M., y Gerdtham, U. (2004). Absolute Income, Relative Income, Income Inequality, and Mortality. *The Journal of Human Resources*, *39*(1), 228–247.

Johnson, E., Rymon, T., y Sen, S. (2003). The Small-Sample Bias of the Gini Coefficient : Results and Implications for Empirical. *The Review of Economics and Statistics*, *85*(1), 226–234.

Kakwani, N. C. (1974). A Note on the Efficient Estimation of the New Measures of Income Inequality. *Econometrica*, *42*(3), 597–600.

Kakwani, N. C., y Podder, N. (2008). Efficient Estimation of the Lorenz Curve and Associated Inequality Measures from Grouped Observations. *Econometrica*, *44*(1), 137–148. https://doi.org/10.1007/978-0-387-72796-7_4

Kokko, H., Mackenzie, A., Reynolds, J. D., Lindström, J., y Sutherland, W. J. (1999). Measures of Inequality Are Not Equal. *The American Naturalist*, *154*(3), 358–382. https://doi.org/10.1086/303235

Krieckhaus, J., Son, B., Bellinger, N. M., y Wells, J. M. (2013). Economic inequality and democratic support. *Journal of Politics*, *76*(1), 139–151. https://doi.org/10.1017/S0022381613001229

Krishnan, P. (1981). Measures of Inequality for Qualitative Variables and Concentration Curves. *American Sociological Review*, *46*(3), 368–371.

Lang, V. F., y Lingnau, H. (2015). Defining and Measuring Poverty and Inequality Post-2015. *Journal of International Development*, *27*(2), 399–414. https://doi.org/10.1002/jid

Leigh, A. (2007). How Closely Do Top Income Shares Track Other Measures of Inequality? *The Economic Journal*, 117(524), F619–F633. https://doi.org/10.1111/j.1468-0297.2007.02099.

Lerman, R. I., y Yitzhaki, S. (1985). Income Inequality Effects by Income Source : A New Approach and Applications to the United States. *The Review of Economics and Statistics*, *67*(1), 151–156.

Li, H., Xie, D., y Zou, H.-F. (2000). Dynamics of Income Distribution. *The Canadian Journal of Economics*, *33*(4), 937–961.

McDonald, J. B., y Jensen, B. C. (1979). An analysis of some properties of alternative measures of income inequality based on the gamma distribution function. *Journal of the American Statistical Association*, *74*(368), 856–860. https://doi.org/10.1080/01621459.1979.10481042

McGranahan, D. (1981). The Meaning and Measurement of Income Inequality. *American Sociological Review*, *46*(2), 240–241.

Mehran, F. (1976). Linear Measures of Income Inequality. *Econometrica*, *44*(4), 805–809.

Menard, S. (1986). A Research Note on International Comparisons of Inequality of Income. *Social Forces*, *64*(3), 778–793.

Mills, J. A., y Zandvakili, S. (1997). Statistical Inference Via Bootstrapping for Measures of Inequality. *Journal of Applied Econometrics*, *12*(2), 133–150.

Minarik, J. J. (1977). The Measurement and Trend of Inequality : Comment. *The American Economic Review*, *67*(3), 513–516.

Formby, J., y G. Seaks, T. (1980). Paglin's Gini Measure of Inequality : A Modification. *The American Economic Review*, *70*(3), 479–482.

Paglin, M. (1975). American Economic Association The Measurement and Trend of Inequality : A Basic Revision. *The American Economic Review*, *65*(4), 598–609.

Paglin, M. (1989). On the Measurement and Trend of Inequality : Reply. *The American Economic Review*, *79*(1), 265–266.

Perotti, R. (1992). Income Distribution, Politics, and Growth. *The American Economic Review*, *82*(2), 311–316.

Praag, B., Hagenaars, A., y Eck, W. (1983). The Influence of Classification and Observation Errors on the Measurement of Income Inequality. *Econometrica*, *51*(4), 1093–1108.

Pyatt, G. (1987). Measuring Welfare, Poverty and Inequality. *The Economic Journal*, *97*(386), 459–467.

Reardon, S. F., y Bischoff, K. (2011). Income Inequality and Income Segregation. *American Journal of Sociology*, *116*(4), 1092–1153. https://doi.org/10.1086/657114

Sastry, D., y Kelkar, U. R. (1994). Note on the Decomposition of Gini Inequality. *The Review of Economics and Statistics*, *76*(3), 584–586.

Schwartz, J., y Winship, C. (1980). The Welfare Approach to Measuring Inequality. *Sociological Methodology*, *11*(1), 1–36.

Smeeding, T. M. (2009). New Comparative Measures of Income, Material Deprivation, and Well-Being. *Journal of Policy Analysis and Management*, *28*(4), 745–752. https://doi.org/10.1002/pam

Thomas Demuynck, y Gaer, D. Van de. (2012). Inequality Adjusted Income Growth. *Economica*, *79*(316), 747–765. https://doi.org/10.1111/j.

Thurow, L. C. (1971). The Income Distribution as a Pure Public Good. *The Quarterly Journal of Economics*, *85*(2), 327–336.

Tillé, Y., y Langel, M. (2012). Histogram-based interpolation of the Lorenz curve and Gini index for grouped data. *American Statistician*, *66*(4), 225–231. https://doi.org/10.1080/00031305.2012.734197

Tinbergen, J. (2018). American Economic Association Welfare Economics and Income Distribution. *The American Economic Review*, *47*(2), 490–503.

World Bank Group. (2016). *Poverty and Shared Prosperity 2016*. *World Bank GRoup*.

World Bank Group. (2016). *Poverty and Shared Prosperity: Taking on Inequality*. (World Bank Group, Ed.) (First Edition). Washington, DC: The World Bank. https://doi.org/10.1596/978-1-4648-0958-3

Yntema, B. Y. D. B. (1933). Measures of the Inequality in the Personal Distribution of Wealth or Income. *Journal of the American Statistical Association*, *28*(184), 423–433.