ESTUDIOS DE ECONOMÍA APLICADA N° 19, 2001. PÁGS. 21-36

Un análisis de la desigualdad de la renta entre las provincias andaluzas, a partir de los datos de la E.B.P.F. y de renta corregidos (1990-91), utilizando estimaciones de combinaciones lineales convexas de curvas de Lorenz

[*]HERRERÍAS PLEGUEZUELO, R. y [**]GARCÍA FERNÁNDEZ, R. Mª.
Departamento de Economía Aplicada. Facultad de Ciencias Económicas y Empresariales. Universidad de Granada.
Telf.: 958 24 37 21-Fax: 958 24 40 46 • [*]e-mail: rherreri@platon-ugr.es - [**]e-mail: rosamgf@ugr.es

RESUMEN

En este trabajo, en primer lugar, se estudian determinadas especificaciones de la curva de Lorenz que cumplen simultáneamente el teorema de Casas y Núñez (1987), se derivan del concepto de función generadora de curva de Lorenz (Callejón, 1995) y pueden estimarse por mínimos cuadrados ordinarios.

En segundo lugar, se completa la demostración de que cualquier combinación lineal convexa de n curvas de Lorenz, cumple el teorema de Casas y Núñez (1987).

En tercer lugar, se han estimado por mínimos cuadrados ordinarios, aquellas combinaciones lineales convexas de curvas de Lorenz que cumplen las anteriores condiciones. Estas estimaciones se han realizado para Andalucía, para lo cual se han utilizado los datos suministrados por la E.B.P.F. 1990-91 y los datos de renta corregidos según la tasa de ocultación progresiva determinada por Pena y otros (1996).

Tras seleccionar aquella combinación lineal en la que la diferencia entre el índice de Gini empírico y estimado es menor, tanto con los datos de la E.B.P.F., como con los datos de renta corregidos, se han realizado las dieciséis estimaciones provinciales. Finalmente, en función de los índices de Gini se ha estudiado la desigualdad de la concentración de la renta en las provincias andaluzas y en función de ello, se ha establecido un orden entre éstas.

Palabras clave: Desigualdad de renta, funciones generadoras, estimación de curvas de Lorenz, combinaciones lineales convexas de curvas de Lorenz, índice de Gini.

ABSTRACT

In this paper, firstly, different functional forms of the Lorenz curve are studied, which satisfy Casas and Núñez's (1987) theorem, they can be obtained using Lorenz curve generator function and finally can be estimated by the method of the least square.

Secondly, it is demonstrated that all lineal convex combinations of *n* Lorenz curves verify the Casas and Nuñez's (1987) theorem.

Código UNESCO: 530204
Artículo recibido el ????? ??????? ????????? ???????

Thirdly, the combinations of Lorenz curves which satisfy the previous conditions have been estimated by least square. The Lorenz curve have been estimated using data from Andalusia. These data have been provided by the E.B.P.F. 1990-91 and Pena at al. work (1996).

After selecting the lineal combination in which the difference between the Gini empirical index and the estimated one is the smallest, for both types of data, have been estimated the Lorenz curve for provinces of Andalusia. Finally using the Gini index has been studied the concentration of the income, and with this information the provinces have been ordered.

Key Words: Income inequality, Lorenz curves generator function, lineal convex combination of Lorenz curves, Gini index.

1. Introducción

El presente trabajo toma en consideración un estudio de Herrerías y García (2000) realizado sólo para el ámbito geográfico de la provincia de Granada, aunque con una orientación diferente.

Se han considerado las especificaciones de curva de Lorenz que cumplen las siguientes condiciones:

i) La condición necesaria propuesta por Casas y Núñez (1987).
ii) Que pueden obtenerse a partir del concepto de función generadora (Callejón, 1995).
iii) Que pueden estimarse por mínimos cuadrados ordinarios.

Por otra parte, Casas, Herrerías y Núñez (1990) demostraron que la combinación lineal convexa formada por las curvas de Lorenz potencial y exponencial satisface la condición necesaria de Casas y Núñez (1987). En este mismo trabajo, estos autores demostraron que cualquier combinación lineal convexa formada por n curvas de Lorenz cumple las cuatro primeras condiciones de dicha condición necesaria, dependiendo el cumplimiento de las dos últimas condiciones de las formas funcionales seleccionadas. En este trabajo se generaliza este último resultado, y se demuestra que cualquier combinación lineal convexa formada por n curvas de Lorenz, cumple la condición necesaria enunciada en i).

Tras realizar las estimaciones de curvas de Lorenz individuales para el conjunto de la comunidad andaluza, se plantean diversas combinaciones lineales de curvas de Lorenz, considerándose obviamente, sólo aquellas cuyas estimaciones de los parámetros de la combinación lineal sean positivos.

La variable utilizada ha sido la renta disponible per capita derivada de la E.B.P.F.90-91. Debido al nivel de ocultación que presentan los datos (Pena y otros, 1996), se ha considerado conveniente plantear otras estimaciones utilizando los datos de renta corregidos.

Para finalizar, se han calculado los índices de Gini empíricos (para datos corregidos y sin corregir) y estimados para Andalucía, asi como la diferencia entre ambos, observándose que esta diferencia es menor cuando se trabaja con combinaciones lineales convexas.

En función de estos resultados se ha estimado la combinación lineal convexa que presenta menor diferencia entre el índice de Gini empírico y estimado a nivel provincial (combinación formada por la segunda especificación de Kalwani y Podder (1973) y Lafuente (1992)) y seguidamente se calculan los índices de Gini correspondientes a cada una de las provincias andaluzas. Utilizando éstos se estudia la desigualdad y se establece una ordenación entre las distintas provincias andaluzas.

2. Consideraciones previas

Desde que Max Otto Lorenz introdujera en 1905 un método gráfico para representar y analizar el tamaño de distribución de renta, se han propuesto diversos métodos de estimación de la curva de Lorenz.

Kakwani y Podder (1973) obtuvieron una condición necesaria que debe satisfacer toda curva de Lorenz, es decir, si $q = q(p)$, representa una curva de Lorenz debe cumplir las siguientes propiedades:

a) $p = 0, \ q(p) = 0$
b) $p = 1, \ q(p) = 1$
c) $q(p) < p$
d) $q(p)$ debe ser monótona creciente.

Casas y Núñez (1987) propusieron el siguiente teorema que constituye una condición necesaria que deben cumplir aquellas especificaciones que se utilicen para estimar curvas de Lorenz.

Teorema 1

Si la variable aleatoria subyacente X es de tipo continuo y no negativa, entonces la función $q = q(p)$ que define la curva de Lorenz verifica las siguientes propiedades:

a) $q(0) = 0$
b) $q(1) = 1$
c) $q'(p) \geq 0 , \ 0 \leq p < 1$
d) $q''(p) \geq 0 , 0 < p < 1$
e) $q \leq p , 0 \leq p \leq 1$
f) $\int_0^1 q(p)dp \leq 1/2$

Esta serie de propiedades hacen posible obtener una gama coherente de formas funcionales para poder estimar curvas de Lorenz, a partir de una distribución de frecuencias observadas.

Del conjunto de curvas de Lorenz que son estimables por mínimos cuadrados ordinarios se han seleccionado las siguientes especificaciones:

-Primera especificación de Kakwani y Podder (1973):

$$L(p) = pe^{-\tilde{a}(1-p)}, \quad > \tag{1}$$

-Segunda especificación de Kakwani y Podder (1973) :

$$L(p) = p^{b}e^{-\tilde{a}(1-p)}, \tilde{a} > 0 \tag{2}$$

-Especificación de Gupta (1984) :

$$L(p) = pA^{p-1}, \quad A>1 \tag{3}$$

-Especificación de Basmann, Hayes, Slottje y Johnson (1990):

$$L(p) = p^{ap+b}e^{-g(1-p^{2})-h(1-p)}, \quad 0 \le p \le 1 \tag{4}$$

Esta especificación sólo verifica el anterior teorema para determinadas especificaciones de sus parámetros. Casas, Herrerías y Núñez (1990) demostraron que para a=0, la expresión siguiente:

$$L(p) = p^{b}e^{-g(1-p^{2})-h(1-p)}, b \ge 1, g \ge 0, h \ge 0 \tag{5}$$

satisface la condición necesaria establecida en el Teorema 1.

Como caso particular de (5), para g = 1 y h = 0, se obtiene la especificación de curva de Lorenz propuesta por Lafuente (1995):

$$L(p) = p^{b}e^{p^{2}-1}, \quad 0 < b \le 1, \; b = 1 \tag{6}$$

De igual modo, para g = 0 y h = 0, se obtiene la curva de Lorenz propuesta por Casas y Núñez (1991):

$$L(p) = p^{b}, b \ge 1 \tag{7}$$

La primera y la segunda especificación de curvas de Lorenz dadas por Kakwani y Podder también se pueden obtener a partir de la expresión (5), haciendo b = 1, g = 0 y h = γ se obtiene la primera especificación y haciendo g = 0 , h = γ se obtiene la segunda especificación.

3. Funciones generadoras de curva de Lorenz

En 1995, Callejón propuso la utilización del concepto de función generadora para modelizar curvas de Lorenz. De esta forma, si la modelización L(p) de una curva de Lorenz verifica la siguiente ecuación diferencial:

$$g(p) = \frac{f'(p)}{f(p)} \qquad \textbf{(8)}$$

ha de ser de la forma: $L(p) = Ce^{\int^{p} g(p)dp}$ **(9)**

sujeta a las condiciones:

a) $\lim\limits_{p \to 0} L(p) = 0$ b) $L(1) = 1$

c) $L'(p) > 0, \ \forall p \in [0,1]$ d) $L''(p) > 0, \ \forall p \in [0,1]$

En el trabajo reseñado anteriormente se estudiaron las condiciones que debe cumplir la función generadora, g(p), para que sea una función generadora de curva de Lorenz.

Estas condiciones son las siguientes:

1) $\lim\limits_{p \to \infty} G(p) = -\infty$, siendo $G(p) = \int^{p} g(p)dp$

2) $g(p) > 0 \quad \forall\, p \in [\,0,1\,]$

3) $(g(p))^2 + g'(p) \geq 0 \quad \forall\, p \in [\,0,1\,]$

En el siguiente cuadro se muestran distintas funciones generadoras y la especificación de curva de Lorenz que modelizan:

Cuadro 1: Funciones generadoras y curvas de Lorenz que modelizan.

Función Generadora de Curva de Lorenz	**Curvas de Lorenz**	
$g(p) = (1 + ã p)/p; ã > 0$	$L(p) = pe^{-ã(1-p)}, \ >$	**(1)**
$g(p) = \frac{b}{p} + ã$	$L(p) = p^{b}e^{-ã(1-p)}, \ >$	**(2)**
$g(p) = \frac{b}{p}$	$L(p) = p^{b}; b \geq 1$	**(7)**
$g(p) = 2p + \frac{b}{p}$	$L(p) = p^{b}e^{p^{2}-1} \quad 0 \leq p \leq 1, b \geq 1$	**(6)**
$g(p) = \frac{2p^{2}g + (aLnp + a + h)p + b}{p}$	$L(p) = p^{ap+b}e^{-g(1-p^{2}) - h(1-p)}$	**(4)**
$g(p) = \frac{1}{p} + LnA$	$L(p) = pA^{p-1}$	**(3)**

Como se observa todas las curvas de Lorenz seleccionadas pueden ser tratadas bajo el supuesto de funciones generadoras, Callejón (1995).

4. Estimación de la curva de Lorenz

De las especificaciones de curva de Lorenz que aparecen en el segundo epígrafe, se han estimado aquellas que verifican la ecuación (8) y permiten utilizar el método de los mínimos cuadrados ordinarios, es decir, las dos especificaciones propuestas por Kakwani y Podder (1973), la especificación de Casas y Núñez (1991) y la especificación de Lafuente (1995). La curva de Lorenz propuesta por Gupta (1984) se omitirá, ya que al linealizar su expresión para aplicar mínimos cuadrados ordinarios se observa que coincide con la primera especificación de Kakwani y Podder (Herrerías R., García, RM (2000)).

Tampoco se incluye la estimación de la especificación de Basmann y otros (1990), porque los parámetros estimados, con los datos de la E.B.P.F. 90-91 y datos de renta corregidos disponibles, no cumplen la condición necesaria del Teorema 1.

Para estimar las cuatro especificaciones de curva de Lorenz que cumplen las condiciones señaladas en la introducción, propuestas por Kakwani y Podder (1973), Casas y Núñez (1991) y Lafuente (1995) se ha utilizado la variable Renta Disponible per capita, suministrada por la E.B.P.F. (90-91).

Como se pone de manifiesto en el trabajo de Pena y otros (1996), si se comparan los datos de renta derivados de la E.B.P.F. (90-91) con los datos de renta procedentes de la Contabilidad Nacional, se observa una ocultación en la declaración de ingresos que se agudiza en familias con mayor nivel de ingresos.

Por este motivo se ha considerado conveniente, estimar adicionalmente las curvas de Lorenz asociadas a la Renta Disponible per capita corregida según una tasa de ocultación progresiva (Pena y otros, 1996).

Tras ordenar los registros del fichero de trabajo[1] de forma creciente, se ha calculado para cada percentil la proporción acumulada de renta disponible per capita. Las estimaciones obtenidas para la comunidad andaluza (siendo el tamaño muestral de 3674 hogares) son las siguientes:

1. Cedidos gentilmente por el Profesor Núñez, de la Universidad de Alcalá de Henares, al cual quedamos agradecidos.

Cuadro 2: Curvas de Lorenz estimadas.

	Datos sin corregir	**Datos corregidos**
Kakwani y Podder (1)	$L(p) = pe^{-1.330(1-p)}$	$L(p) = pe^{-1.414(1-p)}$
Kakwani y Podder (2)	$L(p) = p^{1.342}e^{-0.581(1-p)}$	$L(p) = p^{1.332}e^{-0.705(1-p)}$
Casas y Núñez	$L(p) = p^{1.567}$	$L(p) = p^{1.615}$
Lafuente	$L(p) = p^{1.0968}e^{p^2-1}$	$L(p) = p^{1.1363}e^{p^2-1}$

5. Combinaciones lineales convexas

Casas, Herrerías y Núñez (1990) demuestran en primer lugar, que la combinación lineal convexa formada con las curvas de Lorenz potencial y exponencial, es otra curva que verifica las seis condiciones del Teorema 1 anterior, que toda curva de Lorenz debe cumplir. En segundo lugar, que cualquier combinación lineal convexa compuesta por n curvas de Lorenz, cumple las cuatro primeras condiciones necesarias del mencionado teorema, dependiendo de las formas funcionales seleccionadas el cumplimiento de las dos últimas condiciones. Veamos que, en cualquier caso, estas dos condiciones también las satisface la combinación lineal convexa (resultado conocido y utilizado por Casas y Núñez (1991), no estando publicada su demostración):

La quinta condición necesaria del Teorema 1 requiere que se verifique:

$$\sum_{i=1}^{n} c_i q_i(p) \leq p$$

Se demuestra que la anterior condición se cumple independientemente de la forma seleccionada:

$$\sum_{i=1}^{n} c_i q_i(p) \leq \sum_{i=1}^{n} c_i p = p\sum_{i=1}^{n} c_i = p \quad \text{c.q.d.}$$

La sexta condición necesaria, es inmediata si se verifica la quinta condición.

En función de la anterior demostración puede enunciarse el siguiente Teorema:

Teorema 2

Cualquier combinación lineal convexa formada con n curvas de Lorenz verifica las seis condiciones del Teorema 1.

DEMOSTRACIÓN:

Las cuatro primeras condiciones pueden verse en Casas, Herrerías y Núñez (1991) y las últimas en el párrafo precedente, c.q.d.

De acuerdo con el anterior teorema, se han estimado combinaciones lineales convexas de dos curvas de Lorenz [2], es decir:

$$L(p) = C_1 L_1(p) + C_2 L_2(p); \quad C_1 \geq 0 \;,\; C_2 \geq 0 \;,\; C_1 + C_2 = 1$$

donde $L_1(p)$ y $L_2(p)$ son diferentes especificaciones de curva de Lorenz, seleccionadas entre las cuatro consideradas en el cuadro 2.

Para estimar la anterior expresión se sustituyen los valores de y por sus respectivas estimaciones, de forma que:

$$\hat{L}(p) = C_1 \hat{L}_1(p) + C_2 \hat{L}_2(p) \tag{10}$$

posteriormente se calculan los valores de C_1 y C_2 que hacen mínima la suma de cuadrados de los residuos, es decir, se miminiza la expresión[3]:

$$\text{Min} \sum_{t=1}^{n-1} e_t^2 = \sum_{t=1}^{n-1} \left\{ q_t - \left[C_1 \hat{L}_1(p_t) + (1 - C_1) \hat{L}_2(p_t) \right] \right\}^2 \tag{11}$$

obteniéndose:

$$C_1 = \frac{\sum_{t=1}^{n-1} q_t (\hat{L}_1(p_t) - \hat{L}_2(p_t)) - \sum_{t=1}^{n-1} (\hat{L}_1(p_t) - \hat{L}_2(p_t)) \hat{L}_2(p_t)}{\sum_{t=1}^{n-1} (\hat{L}_1(p_t) - \hat{L}_2(p_t))^2} \quad \text{y} \quad C_2 = 1 - C_1 \tag{12}$$

2. El utilizar sólo dos curvas de Lorenz para formar la combinación lineal convexa, es debido a la complejidad de las expresiones de los parámetros de la combinación, cuando el número de curvas es mayor que dos.

3. es una función cuadrática en p, por lo que sólo presenta un mínimo o un máximo.

De las posibles combinaciones convexas que se obtienen con las cuatro curvas seleccionadas, sólo son factibles las cuatro siguientes:

Si L_1 se corresponde con la primera especificación propuesta por Kakwani y Podder (1973) y L_2 es la especificación de Casas y Núñez (1991), se obtiene:

$$\hat{L}_1(p)=0.8980\hat{L}_{KP1}(p)+0.1020\hat{L}_{CN}(p) \tag{13}$$

Si se realiza la combinación lineal para las dos especificaciones de Kakwani y Podder (1973) se obtiene:

$$\hat{L}_2(p)=0.8174\hat{L}_{KP1}(p)+0.1826\hat{L}_{KP2}(p) \tag{14}$$

Para las curvas de Lorenz propuestas por Kakwani y Podder (1973, segunda especificación) y Lafuente (1995) se tiene:

$$\hat{L}_3(p)=0.6067\hat{L}_{KP2}(p)+0.3933\hat{L}_{Lf}(p) \tag{15}$$

Por último se ha estimado esta combinación lineal convexa para las curvas de Lafuente (1995) y Casas y Núñez (1991):

$$\hat{L}_4(p)=0.5721\hat{L}_{Lf}(p)+0.4279\hat{L}_{CN}(p) \tag{16}$$

Las dos restantes combinaciones se han excluido por ser $C_1 \leq 0$.

En la siguiente tabla, se muestran la suma de cuadrados de los residuos de las anteriores estimaciones y de cada una de las especificaciones utilizadas:

Tabla 1: Suma de cuadrados de residuos.

SSEKP1	SSEKP2	SSECN	SSELf	SSE L1	SSE L2	SSE L3	SSE L4
0.047253	0.11250	0.3748	0.22122	0.04325	0.04411	0.03368	0.02616

Se observa que las combinaciones lineales convexas estimadas, tienen menor suma de cuadrados de residuos que las estimaciones de curvas individuales, siendo la combinación de las curvas de Lafuente (1995) y Casas y Núñez (1991) la que menor suma de cuadrados residual presenta.

Se han replicado los cálculos utilizando datos de renta corregidos, pero en este caso el valor de C_1 correspondiente a la segunda combinación lineal propuesta anteriormente (14) no es positivo. Se comprueba sin embargo, que la combinación de las curvas propuestas por Kakwani y Podder (1973, primera especificación) y Lafuente (1995) es factible, lo cual no ocurría para datos de renta sin corregir.

Los resultados obtenidos son los siguientes:

$$\hat{L}_1(p)=0.007497\hat{L}_{CN}(p)+0.9925\hat{L}_{KP1}(p) \quad \textbf{(17)}$$

$$\hat{L}_2(p)=0.8416\hat{L}_{KP1}(p)+0.1585\hat{L}_{Lf}(p) \quad \textbf{(18)}$$

$$\hat{L}_3(p)=0.48601\hat{L}_{KP2}(p)+0.5140\hat{L}_{Lf}(p) \quad \textbf{(19)}$$

$$\hat{L}_4(p)=0.6805\hat{L}_{Lf}(p)+0.3195\hat{L}_{CN}(p) \quad \textbf{(20)}$$

La tabla siguiente muestra la suma de cuadrados residuales para datos de renta corregidos:

Tabla 2: Suma de cuadrados de residuos para datos corregidos.

SSEKP1	SSEKP2	SSECN	SSELf	SSE L1	SSE L2	SSEL3	SSEL4
0.06501	0.1576	0.5181	0.1461	0.06498	0.06203	0.04919	0.04083

Como puede observarse, se mantienen las conclusiones referentes a datos de renta sin corregir.

Finalmente destacar que tanto para datos corregidos como para datos sin corregir:

– La primera especificación de Kakwani y Podder (1973) es la que mejor se adecua individualmente.

– Las combinaciones lineales convexas propuestas tienen menor suma de cuadrados de residuos que las estimaciones individuales.

– La especificación que presenta menor suma de cuadrados de residuos es la combinación lineal convexa formada por las estimaciones de las curvas de Lorenz propuestas por Casas y Núñez (1991) y Lafuente (1995).

6. Estudio de la desigualdad. Índice de Gini

En primer lugar se ha obtenido el índice de Gini empírico y el índice de Gini estimado[4], para Andalucía, correspondiente a cada una de las cuatro estimaciones de curvas de Lorenz seleccionadas y a las distintas combinaciones lineales convexas de curvas de Lorenz estimadas (ANEXO 1).

---

4. El índice de Gini se ha calculado utilizando la expresión: $1-2\int_0^p L(p)dp$, por métodos de integración numérica.

Se observa que la diferencia entre el índice de Gini empírico (IG) y el índice de Gini estimado de una combinación lineal convexa (IGL), es menor que la diferencia entre el índice de Gini empírico y el índice estimado correspondiente a cada una de las curvas que forman la combinación, es decir:

$$\begin{aligned}(IG\text{-}IGL) &\leq (IG\text{-}IGL1)\\ (IG\text{-}IGL) &\leq (IG\text{-}IGL2)\end{aligned} \qquad \mathbf{(21)}$$

siempre que se verifiquen las siguientes condiciones:

i) $C_1 \leq 1$

ii) $\sum_{t=1}^{n-1} L_2(p_t) \geq \sum_{t=1}^{n-1} L_1(p_t)$

La primera condición se verifica siempre, mientras que la segunda requiere que las combinaciones lineales convexas se construyan de forma que se cumpla la condición (ii).

Bajo este resultado cuando se trabaja con combinaciones lineales convexas de curvas de Lorenz no sólo se reduce la suma de cuadrados residual, sino que además se reduce la diferencia entre el índice de Gini empírico y estimado.

En segundo lugar, el disponer de información desagregada por provincias, nos ha permitido realizar estimaciones para cada una de las provincias andaluzas, siendo conscientes de la pérdida de representatividad muestral que supone trabajar a nivel provincial. Para realizar las estimaciones provinciales, se ha seleccionado una combinación lineal convexa de entre las cuatro anteriores. Aunque la combinación que menor suma de cuadrados presenta es la formada por las curvas de Lafuente (1995) y Casas y Núñez (1991), es decir L_4, se ha seleccionado L_3, por ser la que menor diferencia presenta entre el índice de Gini empírico y estimado (ANEXO 1), siendo la diferencia de suma de cuadrados residuales entre ambas combinaciones despreciable (0.00752).

Las estimaciones provinciales, al igual que antes, se han realizado para datos de Renta Disponible per capita corregidos y sin corregir. (ANEXO 2).

Una vez realizadas las estimaciones de la combinación lineal convexa L_4, se han calculado los índices de Gini para Andalucía cada una de sus provincias por métodos de integración numérica. Los resultados obtenidos son:

Tabla 3. Indices de Gini calculados.

	Datos sin corregir		Datos corregidos	
	I.G. EMP.	I.G. EST.	I.G. EMP.	I.G. EST.
Andalucía	0,321601523	0,32140266	0,350000563	0,34984141
Almería	0,343768808	0,34480659	0,371200042	0,37271582
Cádiz	0,332444254	0,33274829	0,355457745	0,35607168
Córdoba	0,310792923	0,30942349	0,323179268	0,32208385
Granada	0,315595441	0,31414714	0,341984318	0,34066356
Huelva	0,276520053	0,27500238	0,298303234	0,29690724
Jaén	0,304000078	0,30437509	0,328224777	0,32839946
Málaga	0,3231445	0,32341142	0,346057973	0,34643465
Sevilla	0,318316824	0,31780206	0,347578556	0,3475534

Se observa que la desigualdad es mayor cuando se utilizan datos de renta corregidos, tanto para Andalucía, como para cada una de sus provincias. Este resultado es coherente, ya que cuando se utilizan datos de renta corregidos se tienen en cuenta aquellos ingresos que la EBPF no contabiliza y que son mayores en familias con ingresos más altos.

En función de los datos de la segunda tabla, se puede establecer la siguiente ordenación provincial de menor a mayor desigualdad:

Datos sin corregir: Huelva, Jaén, Córdoba, Granada, Sevilla, Málaga, Cádiz y Almería.

Datos corregidos: Huelva, Córdoba, Jaén, Granada, Málaga, Sevilla, Cádiz y Almería.

En relación con los anteriores resultados destacan las siguientes observaciones:

1. A nivel provincial, la desigualdad es mayor cuando se utilizan datos de renta corregidos.
2. La ordenación provincial que se obtiene cuando se utiliza el índice de Gini empírico coindice con la ordenación obtenida al utilizar el índice de Gini estimado, tanto para datos de renta corregidos, como para datos de renta sin corregir.
3. Huelva es la provincia con menor desigualdad, ya que es la provincia con menor índice de Gini empírico y estimado (Tabla 3). Este resultado se mantiene tanto si se utilizan datos de renta corregidos, como si se utilizan datos de renta sin corregir.
4. De igual forma las provincias con mayor desigualdad son Almería y Cádiz.

Se han comparado los índices de Gini empíricos (IG) y estimados (IGL3) de cada una de las provincias, con el índice de Gini de Andalucía (IGA). Para esto, se ha hecho cien el índice de Gini de Andalucía. Los resultados obtenidos se recogen en la siguiente tabla:

Tabla 4. Cociente entre los índices de Gini provinciales y el índice de Gini de Andalucia.

	Datos sin corregir		Datos corregidos	
	I.G./IGA	**IGL3/IGAL3**	**IG/IGA**	**IGL3/IGAL3**
Almería	**106,89278**	**107,2818097**	**106,0569849**	**106,5385084**
Cádiz	**103,37148**	**103,5300361**	**101,559193**	**101,7808841**
Córdoba	96,6391329	96,27284665	92,33678571	92,06567341
Granada	98,1324461	97,74254513	97,70964853	97,37656843
Huelva	85,9821964	85,56319353	85,22935899	84,8690954
Jaén	94,5269398	94,7021067	93,77835741	93,87095141
Málaga	**100,479779**	**100,6249979**	98,87354817	99,02619876
Sevilla	98,9786432	98,87972302	99,30799983	99,34598651
Andalucía	100	100	100	100

Se observa que si se utilizan datos de renta sin corregir, las provincias de Almería, Cádiz y Málaga, es decir, las tres con mayor desigualdad en la concentración de los ingresos, tienen unos índices de Gini por encima del índice de Gini de Andalucía. Si se utilizan los datos corregidos, se observa que sólo las provincias de Almería y Cádiz presentan una desigualdad mayor que Andalucía.

8. Conclusiones

- Tras repasar brevemente diferentes especificaciones de curvas de Lorenz que cumplen la condición necesaria de Casas y Núñez (1987), se seleccionan aquellas curvas que cumpliendo esta condición pueden derivarse del concepto de función generadora de curva de Lorenz (Callejón C. J. (1995)) y estimarse por mínimos cuadrados ordinarios. Las curvas seleccionadas son las especificaciones propuestas por Kakwani y Podder (1973), la especificación de Casas y Núñez (1991) y la especificación de Lafuente (1995).
- Se completa la proposición C, de Casas, Herrerías y Núñez (1990), demostrándose que cualquier combinación lineal convexa de n curvas de Lorenz verifica la condición necesaria de Casas y Núñez (1987). Utilizando esta generalización se han estimado distintas combinaciones lineales convexas formadas por dos de las curvas de Lorenz seleccionadas en el epígrafe cuarto y mencionadas anteriormente. Se comprueba que no se pueden estimar las seis combinaciones posibles, ya que en determinadas combinaciones los parámetros estimados de la combinación no son positivos. Adicionalmente, se observa que las combinaciones posibles varían según se trabaje con datos corregidos o con datos sin corregir.

- Las Tablas 1 y 2, muestran que la suma de cuadrados residual es menor si se utilizan combinaciones lineales convexas de curvas de Lorenz. Concretamente, se obtiene que de forma individual es la primera especificación de Kakwani y Podder (1973) la que menor suma de cuadrados residual presenta. En cuanto a las combinaciones de curvas de Lorenz es la cuarta combinación, formada por las curvas de Lafuente (1995) y Casas y Núñez (1991) la que menor suma de cuadrados de residuos tiene. Estos resultados coinciden para datos de renta corregidos y sin corregir.
- Se demuestra que la diferencia entre el índice de Gini empírico y estimado es menor si se utilizan combinaciones lineales convexas de curvas de Lorenz. Por este motivo se ha estimado a nivel provincial la combinación lineal convexa en la que la anterior diferencia es menor, correspondiéndose con la combinación lineal formada por las curvas de Kakwani y Podder (segunda especificación, 1973) y Lafuente (1995), es decir L3. Posteriormente se calcula el índice de Gini empírico y estimado, correspondiente a L3, a nivel provincial de lo que se deduce:
  - La concentración de la renta es mayor para datos de renta corregidos, que para datos de renta sin corregir.
  - Si se ordenan las provincias con respecto a su índice de Gini empírico, se observa que este orden se mantiene para el índice de Gini estimado, para los dos tipos de datos.
  - Huelva es la provincia con menor concentración de renta, siendo Almería y Cádiz las provincias con mayor desigualdad. Estos resultados se observan para datos de renta corregidos y sin corregir.
  - Si se hace cien el índice de Gini de Andalucía, se tiene que para datos de renta sin corregir, los índices de Gini de Almería, Cádiz y Málaga están por encima del índice de Gini de Andalucía, mientras que para datos de renta corregidos, sólo Almería y Cádiz. tienen una desigualdad mayor a la andaluza.

ANEXO 1

Diferencia en valor absoluto entre el índice de Gini empírico y los diferentes índices de Gini estimados

Tabla 5

Datos sin corregir		Datos corregidos		
IG	0,32160152	I IG-IGES I	0,35000056	I IG-IGES I
IGKP1	0,32777246	0,00617094	0,34558451	0,00441605
IGKP2	0,27670849	0,04489303	0,29781656	0,052184
IGLf	0,39034636	0,06874484	0,39903327	0,04903271
IGCN	0,22344504	0,09815648	0,23518636	0,1148142
IGL1	0,3171285	0,00447302	0,34475679	0,00524377
IGL2	0,31844752	0,003154	0,3540533	0,00405274
IGL3	0,32140266	0,00019886	0,34984141	0,00015915
IGL4	0,31892628	0,00267524	0,34667926	0,0033213

ANEXO 2

Estimaciones de la curva de Lorenz obtenida como combinación lineal convexa de las especificaciones de Kakwani y Podder (segunda especificación, 1973) y Lafuente (1995). Entre paréntesis aparece el tamaño muestral correspondiente a cada provincia.

Datos de renta sin corregir:

-Almería (345): $\hat{L}_3(p)=0.5553\hat{L}_{KP2}(p)+0.4447\hat{L}_{Lf}(p)$

-Cádiz (569): $\hat{L}_3(p)=0.6211\hat{L}_{KP2}(p)+0.3790\hat{L}_{Lf}(p)$

-Córdoba(425): $\hat{L}_3(p)=0.6554\hat{L}_{KP2}(p)+0.3446\hat{L}_{Lf}(p)$

-Granada (427): $\hat{L}_3(p)=0.6707\hat{L}_{KP2}(p)+0.3293\hat{L}_{Lf}(p)$

-Huelva (270): $\hat{L}_3(p)=0.7733\hat{L}_{KP2}(p)+0.2267\hat{L}_{Lf}(p)$

-Jaén (440): $\hat{L}_3(p)=0.6307\hat{L}_{KP2}(p)+0.3693\hat{L}_{Lf}(p)$

-Málaga (534): $\hat{L}_3(p)=0.59071\hat{L}_{KP2}(p)+0.4093\hat{L}_{Lf}(p)$

-Sevilla (664): $\hat{L}_3(p)=0.6436\hat{L}_{KP2}(p)+0.3564\hat{L}_{Lf}(p)$

Datos de renta corregidos:

-Almería (354): $\hat{L}_3(p)=0.4467\hat{L}_{KP2}(p)+0.5533\hat{L}_{Lf}(p)$

-Cádiz (569): $\hat{L}_3(p)=0.5220\hat{L}_{KP2}(p)+0.4780\hat{L}_{Lf}(p)$

-Córdoba (425): $\hat{L}_3(p)=0.6170\hat{L}_{KP2}(p)+0.3830\hat{L}_{Lf}(p)$

-Granada (427): $\hat{L}_3(p)=0.5582\hat{L}_{KP2}(p)+0.4418\hat{L}_{Lf}(p)$
-Huelva (270): $\hat{L}_3(p)=0.6994\hat{L}_{KP2}(p)+0.3006\hat{L}_{Lf}(p)$
-Jaén (440): $\hat{L}_3(p)=0.5433\hat{L}_{KP2}(p)+0.4567\hat{L}_{Lf}(p)$
-Málaga (534): $\hat{L}_3(p)=0.5146\hat{L}_{KP2}(p)+0.4854\hat{L}_{Lf}(p)$
-Sevilla (664): $\hat{L}_3(p)=0.5009\hat{L}_{KP2}(p)+0.4991\hat{L}_{Lf}(p)$

Bibliografía

BASMANN, R.L., HAYES K. J., SLOTTJE D.J. y JOHNSON J.D. (1990): "A General functional form for approximating the Lorenz Curve". Journal of Econometrics, 43; 77-90.

CALLEJÓN C., J. (1995): "Funciones generadoras de una curva de Lorenz". Actas IX Reunión ASEPELT-ESPAÑA, 343-350.

CASAS, J.M., HERRERIAS R. y NUÑEZ J.J. (1990): "Familias de formas funcionales para estimar la curva de Lorenz". Actas IV Reunión de ASEPELT-ESPAÑA, 171-176.

CASAS, J.M. y NUÑEZ J.J. (1987): "Algunas consideraciones sobre las medidas de concentración. Aplicaciones". Actas I Reunión de ASEPELT-ESPAÑA, 49-62.

CASAS, J.M. y NUÑEZ J.J. (1991): "Sobre la medición de la desigualdad y conceptos afines". Actas V Reunión ASEPELT-ESPAÑA, 67-73.

I.N.E.(1992): Encuesta de Presupuestos Familiares 1990/91, Metodología.

GUPTA, M.R. (1984): "Functional form for estimating the Lorenz curve". Econometrica, 52 (5): 1313-1314.

HERRERIAS, R. y GARCÍA, R.M. (2000): "Análisis de la Desigualdad de la Renta en Granada, a partir de los datos de la E.P.F. y Diferentes estimaciones de la Curva de Lorenz. Actas XIV Reunión de ASEPELT-ESPAÑA, en CD Rom.

KAKWANI, N.C. (1980): "Functional forms for estimating the Lorenz curve: a reply". Económétrica, 48 (4); 1063-1064.

KAKWANI, N.C. y PODDER,N. (1973): "On the estimation of Lorenz curves from grouped observations". International Economic Review, 14 (2); 278-291.

KAKWANI, N.C. y PODDER, N. (1976): "Efficient estimation of the Lorenz curve and associated inequality measures from grouped observations". Econometrica, 44(1); 137-148.

KENDALL M. y STUART A. (1977): "The Advanced Theory of Statistics". Vol.1, 4ª de. Griffin. London.

LAFUENTE, L.M. (1994): Medidas de cuantificación de la desigualdad de la renta en España según la E.P.F. 1990/91. Tesis Doctoral. Universidad de Murcia

LAFUENTE, L., M. (1995): "Estimación de la curva de Lorenz a través de una nueva forma funcional". Actas IX Reunión ASEPELT-ESPAÑA, 417-423.

PENA, J.B., CALLEALTA, F.J., CASAS, J.M., MEREDIZ, A. y NUÑEZ. J.J. (1996): "Distribución Personal de la Renta en España". Ed. Pirámide. Madrid.