DESIGUALDAD, DIVERSIDAD Y CONVERGENCIA: *(ALGUNOS) INSTRUMENTOS DE MEDIDA*[*]

Francisco J. Goerlich

Correspondencia: Universidad de Valencia. Departamento de Análisis Económico.
Campus de los Naranjos. Av. de los Naranjos, s/n (Ed. Departamental Oriental) 46022 Valencia.
Tel: 96 382 82 46 / Fax: 96 382 82 49
E-mail: Francisco.J.Goerlich@uv.es. Web: http://www.uv.es/-goerlich.

Editor: Instituto Valenciano de Investigaciones Económicas, S.A.
Primera Edición Octubre 1998.
ISBN: 84-482-1936-8
Depósito Legal: V-4088-1998

[*] Este trabajo recoge los aspectos instrumentales de un informe más amplio titulado "**Dinámica de la distribución provincial de la renta. I:** ***Un enfoque desde la óptica de la desigualdad"*** (Goerlich (1998a)) realizado para el Instituto Valenciano de Investigaciones Económicas (I.V.I.E.) y fue utilizado como material en el curso "***Desigualdad, Diversidad y Convergencia: Instrumentos de Medida"*** organizado por dicho instituto. El autor agradece los comentarios realizados a una versión inicial del trabajo a M. Mas, lo que no implica responsabilizarla de los errores que, a buen seguro, todavían subsisten. Se agradece la financiación recibida de la DGICYT, proyecto PB94-1523, y del Instituto Valenciano de Investigaciones Económicas.

DESIGUALDAD, DIVERSIDAD Y CONVERGENCIA:
(ALGUNOS) INSTRUMENTOS DE MEDIDA

Francisco J. Goerlich

R E S U M E N

Este trabajo presenta, en primer lugar, un conjunto amplio de resultados referentes a índices de desigualdad cuando estos se aplican a unidades geográficas que engloban a varios individuos, por ejemplo regiones o países; se ofrece una descripción de dichos índices y de sus propiedades, así como una comparación entre ellos. Se pretende de esta forma ofrecer un instrumental analítico para el análisis, no sólo de la desigualdad entendida en el sentido microeconómico, sino también de fenómenos en los que se pretenda analizar, de forma más general, la diversidad entre unidades económicas, como por ejemplo el debatido problema de la convergencia a nivel macroeconómico. En segundo lugar, se explora el origen de la desigualdad/diversidad examinando que índices son descomponibles en un cierto sentido según el problema que estemos interesados en tratar. El trabajo puede considerarse como un panorama muy selectivo de la literatura, convenientemente adaptado a nuestros propósitos.

PALABRAS CLAVE: Desigualdad, diversidad y convergencia.

A B S T R A C T

This paper offers a wide set of known results on inequality indexes when they are applied to geographical units, such as regions or countries; a description of the indexes and their properties is given, as well as a comparison among them. In this way we offer a set of analytical tools that can be applied, not only for the study of inequality at individual level, but also for the analysis of divergence and convergence among geographical economic units. Eventually, the origins of inequality are explored by examining what indexes can be decomposed into factors according to a given criteria. This paper can be considered as a selective survey of the literature, conveniently adapted to our purposes.

KEY WORDS: Inequality, divergence and convergence.

1. INTRODUCCIÓN.

Este trabajo realiza la exposición de una serie de técnicas con el ánimo de proporcionar un marco de referencia para una mejor comprensión de la evolución dinámica de determinadas variables económicas. Aunque tomaremos como punto de referencia una variable clave en el proceso de crecimiento económico, como es la **renta *per capita***, los instrumentos que expondremos son aplicables mucho más generalmente; en general son aplicables a la medición de la dispersión de cualquier variable, ya sea con contenido económico o no, en particular son especialmente relevantes cuando consideramos la distribución de una variable, por ejemplo la renta, la riqueza o los beneficios, sobre una determinada población, por ejemplo individuos o propietarios del capital. Una vez identificada la evolución dinámica de la variable objeto de estudio, la renta *per capita*, estaremos interesados en conocer los determinantes de dicha evolución, es decir trataremos de examinar, desde diversas perspectivas, que factores han influido en que la dispersión haya aumentado o disminuido, ello nos ayudará a determinar las causas de la desigualdad o a identificar grupos no homogéneos dentro de la población, lo que nos permitirá finalmente obtener conclusiones de política económica.

El presente trabajo se centra exclusivamente en aspectos metodológicos y prácticos, las aplicaciones de las técnicas ilustradas son numerosas y pueden verse, entre otros autores, en Ruiz-Castillo (1987, 1993, 1997), Bosch, Escribano y Sánchez (1989) y Goerlich y Mas (1997, 1998c) a partir de datos microeconómicos, en Cuadrado Roura (1991), Esteban (1994, 1996), Villaverde Castro (1996, 1997) y Goerlich (1998b) a partir de datos regionales y en Goerlich y Mas (1998a,b) a partir de datos de países.

Dos corrientes de literatura que han permanecido separadas, pero que hasta cierto punto son complementarias y cuyas técnicas de análisis pueden combinarse adecuadamente son: (1) la literatura tradicional sobre la desigualdad (Atkinson (1970), Sen (1973), Shorrocks (1980, 1982, 1984), Chakravarty (1990), Esteban y Ray (1993, 1994), Cowell (1995)), centrada fundamentalmente en el estudio de la distribución personal de la renta, y (2) la reciente literatura sobre la convergencia económica (Barro (1991), Barro y Sala-i-Martin (1991,1992,1995), Quah (1993a,b), Sala-i-Martin (1994)), preocupada por la convergencia o divergencia de la renta *per capita* o productividad de diversas unidades geográficas, ya sean regiones o países. Aunque ambas literaturas han tendido a permanecer separadas es evidente que tienen importantes puntos de contacto. Basta para ello ojear los trabajos de Esteban y Ray (1993) o Esteban (1996) sobre la polarización o los de Baumol (1986), DeLong (1988) o Quah (1996a,b,1997) sobre la existencia de clubs de convergencia para darse cuenta de que, a grandes rasgos, se está hablando de conceptos similares, grupos de individuos o regiones que presentan peculiaridades distintas del resto. A este respecto vale la pena señalar que los últimos trabajos de Quah (1996c,1997) incluyen numerosas referencias a la literatura de la desigualdad en un claro intento de tender un puente entre ambas. Así pues aunque la literatura sobre la desigualdad parte del individuo y la del crecimiento de una unidad espacial considerablemente más amplia, las dos tratan de estudiar la evolución en el tiempo de la distribución de una variable económica considerada de especial relevancia desde el punto de vista del bienestar o de la actividad económica. Debe ser obvio que las técnicas de análisis en un tipo de literatura pueden utilizarse satisfactoriamente en el otro, supuesto que disponemos de los datos necesarios. De hecho algunos autores (Rabadán y Salas (1996)) han propuesto medir directamente la convergencia mediante índices de desigualdad; este enfoque, llevado hasta su extremo, podría sufrir de algunas de las críticas de Quah (1993a,b) y Esteban (1996), ya que no parece adecuado reducir el concepto de convergencia a unos pocos estadísticos.

En este trabajo expondremos un conjunto amplio de resultados referentes a índices de desigualdad, suficientemente flexibles como para permitirnos dilucidar que se esconde detrás de la evolución de los mismos. Dado que nuestra unidad de referencia no es necesariamente el individuo, introduciremos la dimensión poblacional en el análisis, de forma que todos los índices analizados serán estadísticos ponderados, donde las ponderaciones vendrán dadas por las frecuencias relativas de la variable objeto de estudio, población relativa en el caso de la renta *per capita*. Esta dimensión poblacional es normalmente recogida por la literatura de la desigualdad pero, sin embargo y sin causa aparente, parece haber sido olvidada por la reciente literatura sobre la convergencia económica.[1]

Finalmente dos breves reflexiones, en primer lugar palabras como desigualdad, diversidad, diferenciación y convergencia son utilizadas como sinónimos en muchas partes del trabajo, esto no deja de constituir, en otros contextos, un cierto abuso del lenguaje; básicamente lo que subyace en todos los índices analizados es una medida de la dispersión de la distribución de una variable y esa dispersión puede ser designada de formas diferentes según en que situaciones. Si la diversidad es buena o mala, si debe aumentarse o disminuirse mediante políticas adecuadas, es algo que depende de juicios de valor y sobre lo que no nos pronunciamos.

[1] No obstante algunos autores si habían observado este olvido. Rabadan y Salas (1996).

En segundo lugar la desigualdad, al igual que el proceso de crecimiento de las economías, es un fenómeno complejo y multidimensional. Por ello todo intento de resumir el proceso de convergencia en un único estadístico está abocado al fracaso. Quah (1993a,b) ha enfatizado satisfactoriamente este punto y a propuesto una serie de instrumentos metodológicos complementarios para analizar la evolución dinámica de distribuciones en el corte transversal (*model of explicit distribution dynamics*), estos instrumentos serán objeto de atención en un trabajo posterior. Este trabajo tan sólo utiliza instrumentos tomados prestados de la literatura de la desigualdad y se estructura de la siguiente forma. La **sección 2** introduce los índices de desigualdad más habituales así como sus propiedades. La **sección 3** explora el origen en la evolución temporal de los índices atendiendo a tres criterios, (i) un criterio basado en particionar la población, (ii) un criterio basado en una desagregación aditiva de la renta y (iii) un criterio basado en una desagregación multiplicativa de la renta *per capita*. Finalmente, la **sección 4** ofrece unas breves conclusiones.

2. DESIGUALDAD Y CONVERGENCIA: ÍNDICES Y PROPIEDADES.

Índices basados en la curva de Lorenz

La literatura de la desigualdad es prolífica en índices (Chakravarty (1990), Cowell (1995)), no entraremos a comentar la mayoría de ellos sino que simplemente utilizaremos algunos de los más populares que se ajustan a nuestras necesidades. En concreto, no expondremos exahustivamente índices relacionados con el enfoque de la función de bienestar social (Atkinson (1970), Sen (1973)), salvo por unos breves comentarios, ni tampoco los recientes indices de polarización (Esteban y Ray (1993, 1994), Esteban (1996)).[2] Al mismo tiempo dejaremos de lado la problemática específica sobre la medición de la desigualdad, para centrarnos en algunos resultados concretos.[3]

Los índices de desigualdad más comunes guardan una relación muy estrecha con la llamada **curva de Lorenz** (Lorenz (1905)) para la distribución de la renta, por ello resulta

[2] El concepto práctico de polarización está todavía en vías de desarrollo, sin embargo la relación entre el concepto de polarización y su íntima relación con la existencia de más de una moda en una distribución hace que sea un concepto muy interesante en el debate de la convergencia *versus* la existencia de clubs.

[3] Zubiri (1985) y Ruiz-Castillo (1986) ofrecen una introducción asequible a dicha problemática. Díaz-Gimenez, Quadrini y Ríos-Rull (1997) y Quadrini y Ríos-Rull (1997) muestran una integración de la literatura tradicional sobre medición de la desigualdad con el moderno análisis teórico de simulación de modelos de equilibrio general capaces de generar desigualdad en un mundo de individuos heterogéneos.

conveniente comenzar nuestra discusión con una exposición de este diagrama. A menos que se indique lo contrario razonaremos en términos de distribuciones discretas, ya que esta es normalmente la naturaleza de los datos disponibles. Supongamos que disponemos de n agrupaciones de individuos[4] cuya **renta *per capita*** designamos por $\boldsymbol{x_i}$, $x_i = Y_i/N_i$, siendo Y_i la renta[5] y N_i la población de la agrupación $i = 1,2,...,n$.[6] Sea además $\boldsymbol{p_i}$ la **frecuencia relativa**, porcentaje de población por agrupación, $p_i = N_i/N$, $N = \Sigma_{i=1}^{n} N_i$; e $\boldsymbol{y_i}$ la proporción de renta de la agrupación i, $y_i = Y_i/Y$, $Y = \Sigma_{i=1}^{n} Y_i$, entonces la **renta *per capita* media** para el agregado puede expresarse como una media aritmética ponderada, $\mu = \dfrac{Y}{N} = \Sigma_{i=1}^{n} p_i x_i$.

Consideremos ahora, sin pérdida de generalidad, que la renta *per capita* de las diversas agrupaciones han sido ordenadas de forma no-decreciente, $x_1 \leq x_2 \leq \ldots \leq x_n$, y llamemos a este vector ordenado de rentas *per capita* $\mathbf{x} = (x_1, x_2, \ldots, x_n)'$. Las agrupaciones se ordenan de acuerdo con **x**. A partir de aquí definimos F_s como la proporción de población que recibe una renta *per capita* igual o inferior a x_s, $F_s = \Sigma_{i=1}^{s} p_i$, y Φ_s como la proporción acumulada de renta correspondiente al nivel de renta *per capita* x_s, $\Phi_s = \Sigma_{i=1}^{s} y_i$, es decir Φ_s es la proporción de renta que reciben las agrupaciones con nivel de renta *per capita* igual o inferior a x_s. Definiendo $F_0 = \Phi_0 = 0$, entonces la relación entre Φ_s y F_s es la llamada **curva de Lorenz**.

La curva de Lorenz, **L(*p*)**, es pues una función de **estadísticos de orden** y muestra el porcentaje acumulado de renta correspondiente al percentil p de la distribución de la renta *per capita*, $0 \leq p \leq 1$. Una representación gráfica, para distribuciones continuas, de la **curva de Lorenz** se muestra en el **gráfico 1**. Por construcción la curva de Lorenz no estará por encima de la recta de 45 grados, si coincide con esta ello implica completa igualdad, y cuanto más lejos esté de dicha recta mayor será el grado de desigualdad observado en la distribución de la renta.

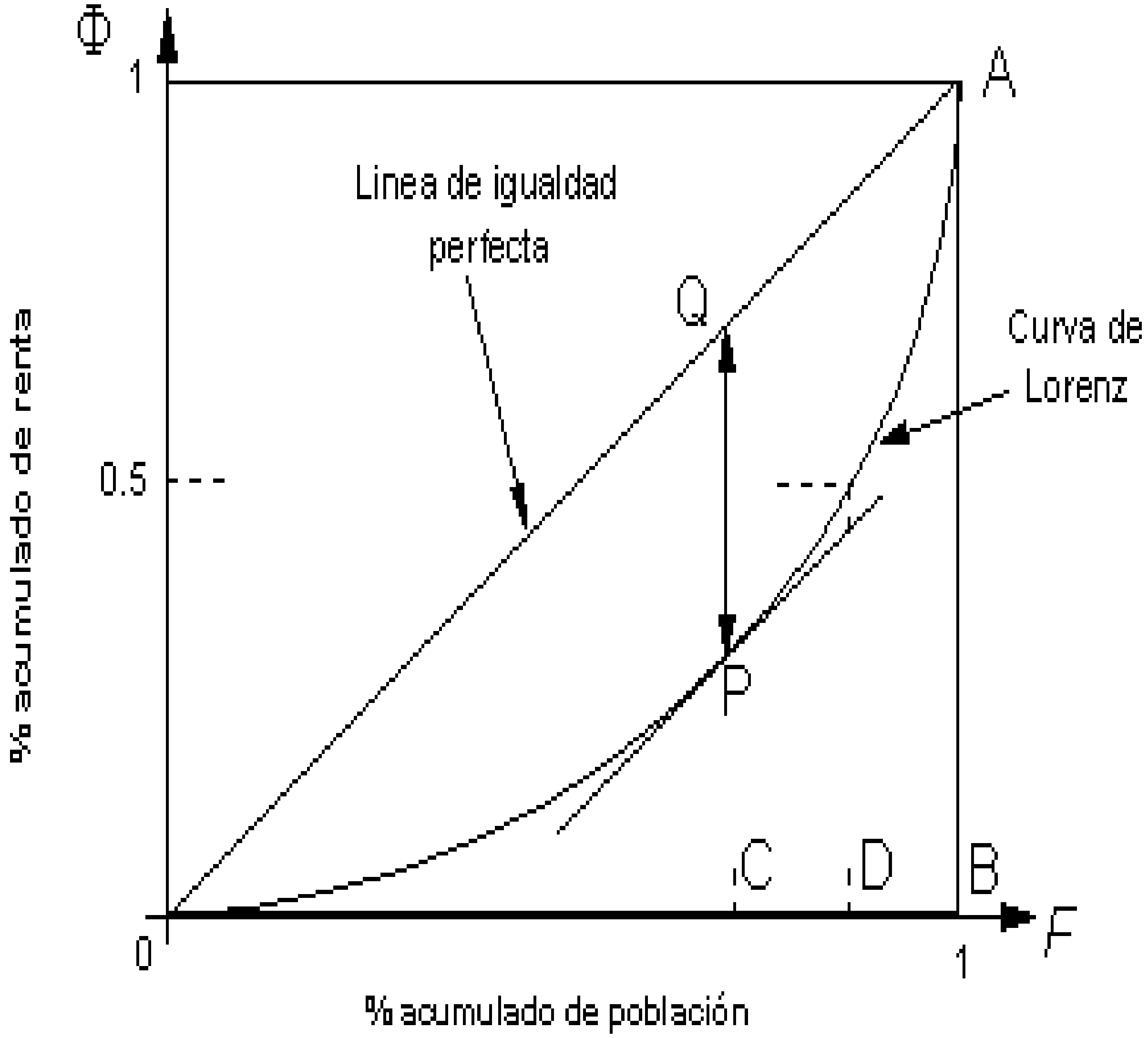
---

[4] La mayor parte de la literatura sobre desigualdad ofrece sus argumentos en términos de individuos u hogares, en otros casos disponemos de observaciones sobre agrupaciones de individuos, por lo que una utilización adecuada de los índices requiere algunas modificaciones en la formulación para aplicar correctamente las ponderaciones poblacionales de cada agrupación. El presente trabajo realiza un esfuerzo en este sentido.

[5] x_i es la **renta real equivalente *per capita***, es decir ha sido adecuadamente deflactada y ajustada por las diferentes necesidades de las agrupaciones, familias o individuos. (Deaton y Muellbauer (1980)).

[6] Aunque es posible estudiar la desigualdad para distribuciones con rentas negativas (Shorrocks (1980), Cowell (1995)) supondremos que $x_i \geq 0 \ \forall i$; muchas medidas estándar de desigualdad no están definidas para rentas negativas y aquellas que lo están necesitan normalmente ser reinterpretadas. Algunas medidas de desigualdad de las que consideraremos no están definidas para distribuciones con rentas nulas.

Gráfico 1: Curva de Lorenz



Es fácil comprobar las siguientes propiedades de la curva de Lorenz:

- $0 \leq F_s \leq 1, 0 \leq \Phi_s \leq 1, \forall s$, $F_0 = \Phi_0 = 0$ y $F_n = \Phi_n = 1$
- L(p) es **convexa** y su derivada, para distribuciones continuas, L′(p), viene dada por $L'(p) = \frac{x}{\mu}$, siendo x el nivel de renta *per capita* correspondiente al percentil p. Para distribuciones discretas L(p) es continua, pero no diferenciable en todos sus puntos, su pendiente es constante entre F_{s-1} y F_s y viene dada por $\frac{\Delta\Phi_s}{\Delta F_s} = \frac{y_s}{p_s} = \frac{x_s}{\mu}$.[7]
- La pendiente de la curva de Lorenz es igual a la unidad en el percentil para el cual el nivel de renta *per capita* es igual a μ, de esta forma la proporción de población que recibe una renta menor o igual a la media puede obtenerse directamente del diagrama de Lorenz. La pendiente de la curva de Lorenz aumenta conforme nos movemos hacia los percentiles superiores, es decir los sectores más ricos, de la población.

La curva de Lorenz correspondiente al caso en que todas las agrupaciones reciben la misma renta *per capita*, $x_i = \mu \; \forall i$, se muestra como la línea 0A en el gráfico 1, esta es la **linea de igualdad perfecta**.[8] Muchos índices intentan resumir la información gráfica suministrada por la curva de Lorenz en una medida cuantitativa que muestre la divergencia entre dicha curva y la situación de igualdad perfecta. La más popular de estas medidas es el **índice de Gini, *G***. (Gini (1912)).[9] Desde el punto de vista geométrico el **índice de Gini** se define como el **cociente del área entre la curva de Lorenz y la linea de igualdad perfecta dividida por el área del triángulo 0AB**, que es igual a ½; por tanto *G* es equivalente a **dos veces el área comprendida entre la curva de Lorenz y la linea de igualdad perfecta**; en consecuencia el índice de Gini varía, para distribuciones continuas, entre 0, igualdad perfecta, y 1, máxima desigualdad, y será mayor cuanto más se aleje la curva de Lorenz de la linea de igualdad perfecta. Alternativamente, y tambien desde un punto de vista geométrico (Rao (1969)), podemos definir el **índice de Gini** como **1 menos dos veces el área bajo la curva de Lorenz**. Sin embargo, desde el punto de vista computacional es conveniente disponer de una fórmula que nos dé el anterior resultado, en tal caso podemos definir el **índice de Gini** como **un medio**

---

[7] Suponemos que los puntos intermedios se obtienen por interpolación lineal, lo cual es el procedimiento habitual. Existen, no obstante, otros procedimientos de interpolación más eficientes (Kakwani y Podder (1973, 1976), Cowell y Mehta (1982)).

[8] En este caso la distribución de la renta *per capita* tiene toda su masa de probabilidad concentrada en un punto, μ.

[9] A pesar de su nombre el índice de Gini (1912) ya fue utilizado por Helmert (1876) y otros estadísticos alemanes en la década de 1870 (David (1968)).

de la diferencia media relativa (Kendall y Stuart (1963))

$$G = \frac{1}{2\mu} \sum_i \sum_j p_i p_j \left| x_i - x_j \right| \qquad (1)$$

donde $\sum_i$ debe entenderse como $\Sigma_{i=1}^n$.

Otra medida alternativa de desigualdad basada en el diagrama de Lorenz viene dada por la **distancia máxima entre la linea de igualdad perfecta y la curva de Lorenz**, ***M***, QP en el gráfico 1 (Schutz (1951)). Dicha distancia se maximiza en el punto donde la curva de Lorenz tiene pendiente igual a 1 y desde el punto de vista estadístico esta medida resulta ser igual a **un medio de la desviación absoluta media relativa** (Sen (1973)),

$$M = \frac{1}{2\mu} \sum_i p_i \left| x_i - \mu \right| \qquad (2)$$

Desde el punto de vista geométrico ***M*** es igual al **cociente del área del mayor triángulo que puede inscribirse entre la curva de Lorenz y la linea de igualdad perfecta dividida por el área del triangulo 0AB**, que es igual a ½. Al igual que el índice de Gini, *M* varía entre 0, igualdad perfecta, y 1, máxima desigualdad, para distribuciones continuas.

Es posible imaginar otras medidas de desigualdad basadas en la curva de Lorenz. Por ejemplo, el porcentaje de población con una renta igual o inferior a μ, 0C en el gráfico 1, o el porcentaje de población para el que $\Phi_k = 0.5$, 0D en el gráfico 1; en ambos casos se trata de **coeficientes** que tratan **de** medir **igualdad entre proporciones**. Estas medidas están sujetas, como veremos a continuación, a la misma crítica que *M* y no son habitualmente utilizadas como indicadores únicos, sino más bien como indicadores complementarios; no obstante el examen de proporciones es muy útil para analizar que es lo que sucede, principalmente, en los extremos de la distribución.

Propiedades básicas

Tres son las propiedades básicas que razonablemente un índice de desigualdad debe satisfacer:

1. Debe ser **independiente de la escala**, es decir, el índice debe permanecer inalterado si la renta de cada individuo en la población (o la renta *per capita* de cada agrupación) se ve alterada en la misma proporción. **Homogeneidad de grado cero en rentas**. Bajo esta propiedad el índice es insensible al nivel de renta medio, lo que implica que la desigualdad es considerada como un problema relativo.[10]

2. Debe ser **independiente del tamaño de la población**, es decir, el índice debe permanecer inalterado si el número de individuos en cada nivel de renta se ve alterado en la misma proporción. **Homogeneidad de grado cero en población**. Con esta propiedad el índice depende sólo de las frecuencias de población relativas en cada nivel de renta, no de las frecuencias de población absolutas.[11]

3. Debe satisfacer el **principio de las transferencias de Pigou (1912)-Dalton (1920)**, esto es, cualquier transferencia de un individuo rico a uno más pobre que no invierta sus *rankings* relativos debe reducir el valor del índice (Sen (1973)).[12]

Las propiedades de independencia respecto a la escala y al tamaño de la población implican, consideradas de forma conjunta, que el índice puede ser calculado a partir de la curva de Lorenz, en otras palabras, bajo estas propiedades la renta media y el volumen total de población son innecesarios para la obtención del índice. Al mismo tiempo un índice que puede ser calculado a partir de la curva de Lorenz obviamente satisface las propiedades de independencia respecto a la escala y al tamaño de la población.

[10] Los índices de desigualdad suelen abstraerse del crecimiento en la renta para concentarse en problemas puramente distributivos. Sin embargo, y aún suponiendo una preferencia por la igualdad, las reducciones a lo largo del tiempo en las desigualdades no deben tomarse como sinónimo de que "la sociedad está mejor ahora que antes"; así por ejemplo ligeras reducciones en la desigualdad acompañadas de importantes caídas en la renta media, y por tanto en el volumen de renta a repartir, pueden ser consideradas como un empeoramiento de la situación. Juicios de valor sobre lo que la evolución de la desigualdad significa a nivel agregado en economías en crecimiento son difíciles de realizar. Algunas modificaciones en las medidas habituales de desigualdad, que tratan de recoger la evolución del volumen total de renta a repartir, han sido propuestas en la literatura (Moyes (1989)), pero no serán objeto de atención en este trabajo. El enfoque de la función de bienestar social, que será analizado más adelante, también puede ser utilizado para analizar esta cuestión (Cowell (1995)).

[11] En el caso de distribuciones continuas esto es equivalente a que el índice pueda ser calculado sólo a partir de la función de densidad.

[12] En la literatura este principio se conoce como el **principio débil de las transferencias** porque todo lo que requiere es que, dada una transferencia, la desigualdad se reduzca; pero no dice nada acerca de cuanto debe reducirse. El **principio de las transferencias en sentido fuerte** requiere que la reducción en la desigualdad dependa sólo de la distancia entre las proporciones de renta de los individuos, y no de los individuos en sí mismos.

Tanto el **índice de Gini**, ***G***, como ***M*** satisfacen las propiedades de independencia respecto a la escala y al tamaño de la población, ya que han sido definidos a partir de la curva de Lorenz. El índice de Gini satisface, además, el principio de las transferencias de Pigou-Dalton; a partir de la definición geométrica del índice de Gini obsérvese que una transferencia de una persona rica a una más pobre eleva toda la curva de Lorenz entre los correspondientes percentiles, y por tanto reduce el valor del índice. Sin embargo aquellas medidas que, como M, son proporcionales a la desviación absoluta media relativa no satisfacen dicho principio; a partir de la definición geométrica de M es claro que dicho índice es invariante a transferencias de renta entre individuos sólo por encima o sólo por debajo de μ; en este caso la distancia QP en el gráfico 1 no se ve alterada.

El principal inconveniente de G es su curiosa valoración de cambios en la distribución de la renta en función de en que parte de la distribución ocurran. En concreto una transferencia de un individuo rico a uno pobre tendrá mucho mayor efecto sobre el índice de Gini si los dos se encuentran cerca del centro de la distribución que si se encuentran en los extremos de la misma. Por otra parte ambos extremos de la distribución son tratados de forma simétrica.

Finalmente podemos preguntarnos que sucede con la desigualdad si la renta *per capita* de cada agrupación se incrementa en la misma cantidad absoluta,[13] digamos $\delta > 0$, en este caso tanto el índice de Gini como M disminuirán como una función de δ. De hecho puede demostrarse con facilidad que $G, M \to 0$ conforme $\delta \to \infty$.

Comparabilidad

Un **índice de desigualdad** es, por tanto, una **representación numérica escalar** de un fenómeno complejo y multidimensional como es la distribución de la renta sobre una población determinada. Lamentablemente es cierto que las comparaciones entre distribuciones son sensibles a la elección del índice de dispersión elegido, ya que las diferentes medidas de desigualdad tienden a enfatizar de forma diferente la desigualdad en distintas partes de la distribución. Resulta por ello útil examinar algunos conceptos relacionados con la comparabilidad entre índices.

Dados dos índices de desigualdad diremos que son **ordinalmente equivalentes** si proporcionan la misma ordenación sobre diferentes vectores de renta, **x**; por tanto ante la

[13] Para una distribución geográfica de la población no uniforme esto no es equivalente a que la renta de cada individuo en la población se vea alterada en la misma cuantía absoluta.

pregunta si la distribución $\mathbf{x}_{(1)}$ es más desigual que la distribución representada por $\mathbf{x}_{(2)}$, todos aquellos índices que sean ordinalmente equivalentes proporcionarán la misma respuesta. En este sentido el concepto de representación numérica asociada al índice hace referencia al concepto de *ranking*.

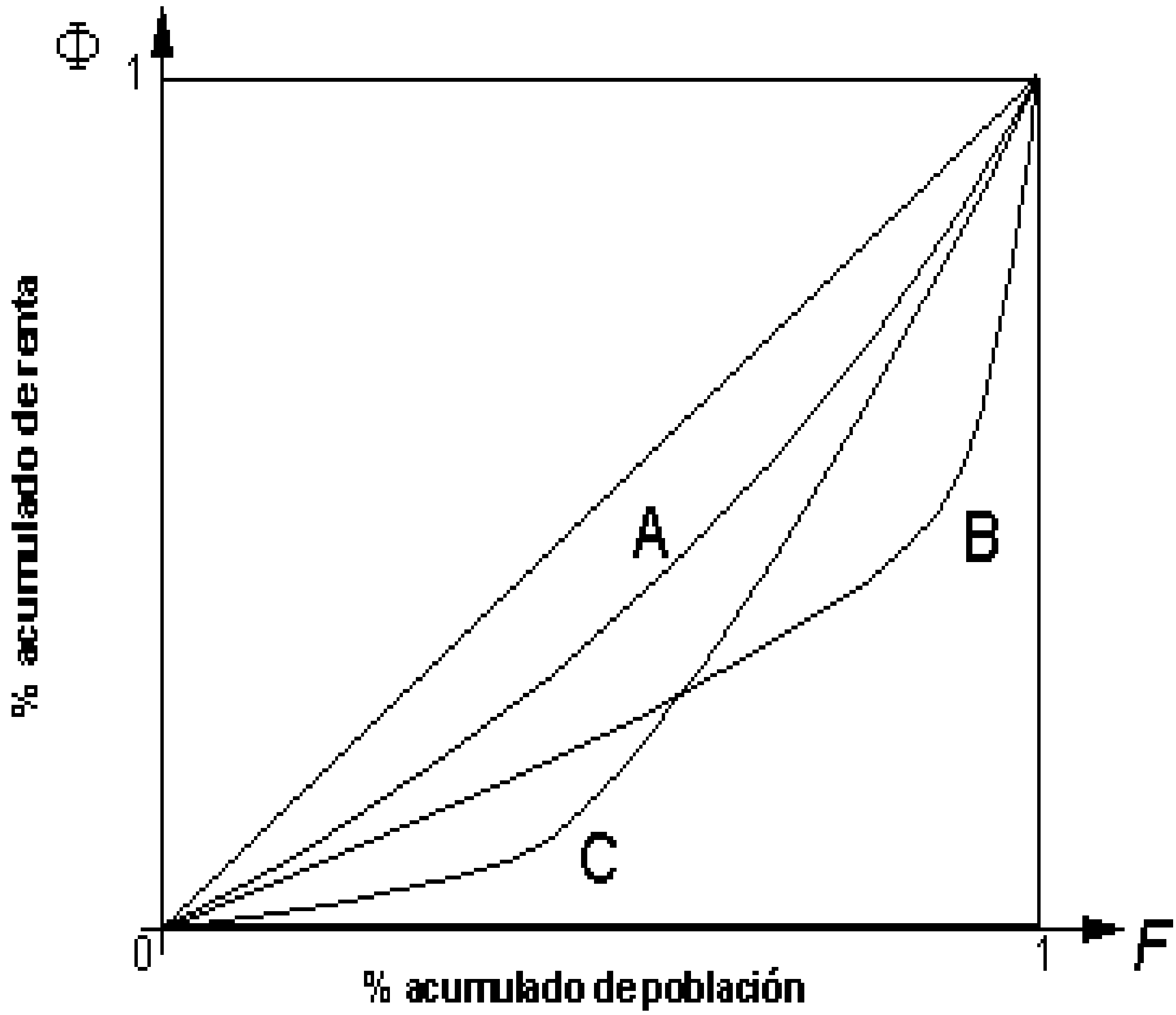
Dos índices de desigualdad diremos que son **cardinalmente equivalentes** si un índice es una transformación lineal del otro; de forma que además de proporcionar la misma ordenación sobre diferentes vectores de renta, $\mathbf{x}$, muestran el mismo porcentaje de variación al comparar dos situaciones diferentes; es decir si comparamos la distribución $\mathbf{x}_{(1)}$ con la representada por $\mathbf{x}_{(2)}$, todos aquellos índices que sean cardinalmente equivalentes nos indicarán el mismo porcentaje de variación en la desigualdad entre ambas situaciones. Obviamente **equivalencia cardinal implica equivalencia ordinal**, pero no a la inversa.

En vista de ello **no** debemos atribuir un gran significado a afirmaciones referentes al porcentaje de reducción de la desigualdad según un determinado índice, a menos que un gran conjunto de índices ordinalmente equivalentes proporcionen los mismos resultados cuantitativos.

El principio de las transferencias de Pigou-Dalton y su relación con los índices de desigualdad puede ser entendido mejor por medio del **gráfico 2** que muestra tres curvas de Lorenz, correspondientes a tres distribuciones diferentes, A, B y C. Puesto que la curva de Lorenz A se sitúa “dentro”, tanto de la curva de Lorenz B como de C, la distribución representada por A puede alcanzarse mediante transferencias de ricos a pobres sin necesidad de que ello altere sus *rankings*, tanto si partimos de la distribución B como de C. En consecuencia cualquier medida de desigualdad que satisfaga el principio de las transferencias de Pigou-Dalton indicará un menor nivel de desigualdad en A, ya sea tanto con respecto a B como con respecto a C, en particular $G_A < G_B$ y $G_A < G_C$. Sin embargo, si comparamos B con C la situación es muy distinta, ambas curvas de Lorenz se intersectan, de forma que ante la pregunta de cual de estas dos distribuciones es más desigual la respuesta dependerá del índice de desigualdad particular que utilizemos. La distribución representada por B es más igual en el extremo inferior frente a la representada por C, que es más igual en el extremo superior. Que clase de desigualdad es “peor” o “mejor” en esta situación es algo que depende de juicios de valor, más allá del principio de las transferencias. No es difícil construir ejemplos en los que $G_B < G_C$ mientras que en otras situaciones $G_C < G_B$ y las curvas de Lorenz se intersectan tal y como se representa en el gráfico 2 (Schutz (1951)). Por ello cuando una distribución **domina** a otra en el **sentido de Lorenz**, tal y como sucede con A respecto a B o C, entonces todas aquellas medidas de desigualdad que satisfagan el principio de las transferencias de Pigou-

Dalton proporcionarán el mismo *ranking* entre distribuciones, pero cuando no se da tal dominancia, como sucede con B y C, el *ranking* dependerá de juicios de valor, implícitos o explícitos, acerca de la importancia de la desigualdad en diferentes partes de la distribución. Es por ello que resultaría útil considerar **familias de índices** dependientes de un parámetro al que atribuir un significado específico en términos de la importancia otorgada a distintas partes de la distribución, de esta forma un índice de desigualdad particular resuelve la ambigüedad producida al comparar situaciones como las representadas por B o C de una forma concreta e interpretable. Los dos epígrafes siguientes se dedican parcialmente a esta cuestión.

Gráfico 2: Curvas de Lorenz con y sin intersección



Índices basados en el concepto de entropía de la teoría de la información

Theil (1967) propuso dos interesantes medidas de desigualdad a partir del concepto de entropía de la teoría de la información.[14] Estos índices pueden obtenerse como casos particulares de la **clase de medidas generalizadas de entropía** (Cowell (1995)) que viene dada por

$$T(\beta) = \frac{1}{\beta(\beta-1)} \sum_i p_i \left[\left(\frac{x_i}{\mu} \right)^{\beta} - 1 \right] \quad \beta \neq 0{,}1 \qquad \text{donde } \mu = \Sigma_{i=1}^{n} p_i x_i \tag{3}$$

y que también es conocida en la literatura como la **familia de índices de Theil**. Mientras que el límite inferior de $T(\beta)$ es siempre 0, igualdad perfecta, el límite superior varía con el valor de β. Para $\beta > 0$ ($\neq 1$) $T(\beta)$ presenta una cota superior, por el contrario cuando $\beta \leq 0$ la familia de índices de Theil no está acotada superiormente.[15]

La familia de índices de Theil cumple la mayoría de propiedades deseables que se le pueden exigir a los índices de desigualdad, es independiente de la escala y del tamaño de la población y satisface el principio de las transferencias de Pigou-Dalton, además es descomponible en un cierto sentido que será analizado más adelante. Esta familia de medidas tiene claras similitudes con las medidas propuestas por Atkinson (1970) a partir del enfoque de la función de bienestar social, tal y como veremos a continuación.

El parámetro β afecta a la sensibilidad del índice ante transferencias entre ricos y pobres en función de en que parte de la distribución se realicen. En este sentido se puede demostrar (Shorrocks (1980)) que, conforme β disminuye el índice $T(\beta)$ es más sensible a transferencias en la parte inferior de la distribución. En el límite, conforme $\beta \to -\infty$, el índice se centra sólo en el extremo inferior de la distribución. Por el contrario, conforme β aumenta el índice $T(\beta)$ se vuelve más sensible a transferencias en la parte superior de la distribución. De hecho para

[14] La **idea básica** asociada a la teoría de la información es la siguiente: Sea *w* la probabilidad de que ocurra un cierto suceso; entonces el contenido informativo de que tal suceso haya ocurrido, *h(w)*, será una función decreciente de *w*, cuanto menos probable sea un suceso, más interesante será conocer que este ha ocurrido. Una función, entre otras posibles, que satisface esta propiedad y esta detrás de la formulación original de los índices de Theil (1967), es $h(w) = \log(1/w)$.

[15] Para $\beta > 0$, $T(\beta)$ puede ser normalizado de forma que su rango de variación se sitúe en el intervalo [0, 1]. Sin embargo esta normalización afecta a la propiedad de descomponibilidad aditiva por subgrupos de población que estudiaremos en la sección siguiente. Para $\beta \leq 0$, $T(\beta)$ no puede ser normalizado sin abandonar la propiedad de descomponibilidad aditiva. (Shorrocks (1980)).

valores de $\beta > 2$ el índice sólo parece mostrar sensibilidad ante la igualación de rentas entre los más ricos de la distribución. Aunque el principio de las transferencias de Pigou-Dalton se satisface para cualquier valor de β, en todo el dominio de definición de rentas, esta característica del índice ha llevado a algunos autores a preguntarse si $T(\beta)$ no debería ser eliminado de su consideración como índice de desigualdad para valores elevados del parámetro β (Kolm (1976a,b), Love y Wolfson (1976)).

Para $\beta = 1$ la regla de L´Hopital aplicada a (3) permite obtener la primera de las medidas propuestas por Theil (1967)

$$T(1) = \sum_i p_i \left(\frac{x_i}{\mu}\right) \log\left(\frac{x_i}{\mu}\right) \tag{4}$$

que varía entre 0, igualdad perfecta, y $-\log p_i$, máxima desigualdad, cuando la agrupación i acapara todo el volumen de renta. Obsérvese que $\frac{x_i}{\mu}$ es simplemente la pendiente de la curva de Lorenz en el percentil correspondiente al nivel de renta *per capita* x_i; por tanto, $T(1)$, al igual que las medidas anteriores, puede ser obtenida directamente a partir de la curva de Lorenz.

Algunas manipulaciones algebraicas permiten entender mejor la estructura del índice. Escribiendo $\frac{p_i x_i}{\mu} = y_i$ observamos que $T(1)$ pondera una medida de desigualdad, $\log\left(\frac{x_i}{\mu}\right)$, por proporciones de renta. Además puesto que $\frac{x_i}{\mu} = \frac{y_i}{p_i}$ podemos escribir $T(1)$ como

$$T(1) = \sum_i y_i \log\left(\frac{y_i}{p_i}\right) \tag{5}$$

lo que permite a Theil (1967, p.-95) interpretar este índice como *"la información esperada de un mensaje que transforma proporciones de población en proporciones de renta"*. En el caso de una igualdad perfecta las proporciones de renta y población de cada agrupación son idénticas y $T(1)$ toma un valor 0.[16] En el caso de mayor desigualdad, cuando hipotéticamente una agrupación dispone de todo el volumen de renta y el resto no disponen de nada, todos los

[16] Puede pensarse en $T(1)$ como una función general de distancia que mide la divergencia entre proporciones de renta y proporciones de población.

elementos del vector **x** son cero, excepto el correspondiente a la agrupación que concentra toda la renta, digamos $x_i = \frac{Y}{N_i}$; es esta situación y_i es igual a 1 e y_j es igual a 0 para $\forall j \neq i$, por tanto $T(1)$ toma su valor máximo, $-\log p_i$, al tender el resto de términos en el sumatorio (5) a 0, puesto que $y.\log y \rightarrow 0$ conforme $y \rightarrow 0$.

Para $\beta = 0$ la regla de L´Hopital aplicada a (3) permite obtener la segunda de las medidas propuestas por Theil (1967)

$$T(0) = -\sum_i p_i \log\left(\frac{x_i}{\mu}\right) \qquad (6)$$

que toma el valor 0 cuando existe igualdad perfecta, pero que sin embargo no está definida para distribuciones con renta *per capita* cero. De nuevo obsérvese que $\frac{x_i}{\mu}$ es simplemente la pendiente de la curva de Lorenz en el percentil correspondiente al nivel de renta *per capita* x_i; por tanto, $T(0)$ puede ser obtenida directamente a partir de la curva de Lorenz.

Esta medida es análoga a la anterior, excepto por el hecho de que intercambia los papeles de las proporciones de renta y población en $T(1)$. Invirtiendo estos términos en (5) obtenemos

$$T(0) = \sum_i p_i \log\left(\frac{p_i}{y_i}\right) = \sum_i p_i \log\left(\frac{\mu}{x_i}\right) \qquad (7)$$

lo que indica que $T(0)$ pondera una medida de desigualdad, $\log\left(\frac{\mu}{x_i}\right)$, por proporciones de población. Theil (1967, p.-125) interpreta este índice como *"el contenido de información esperada de un mensaje indirecto que transforma proporciones de renta como probabilidades a* priori *en proporciones de población como probabilidades a* posteriori*"*. En el caso de una igualdad perfecta las proporciones de renta y población de cada agrupación son idénticas y $T(0)$ toma un valor 0.[17] La expresión (7) para $T(0)$ puede reescribirse como

[17] Al igual que $T(1)$, el índice $T(0)$ trata de medir la divergencia entre proporciones de renta y proporciones de población, pero utiliza unas ponderaciones diferentes sobre la misma función de distancia.

$$T(0) = \log \mu - \sum_i p_i \log x_i = \log \frac{\mu}{\tilde{\mu}} \qquad (8)$$

donde $\tilde{\mu}$ es la media geométrica de la distribución de la renta *per capita*, $\tilde{\mu} = \Pi_{i=1}^{n} x_i^{p_i}$.[18] En otras palabras *T*(0) es el logaritmo de la *ratio* entre la media aritmética y la media geométrica de la distribución de la renta *per capita*.[19] El principal inconveniente de *T*(0) es que no está definido para distribuciones con renta cero, puesto que $\log x \to -\infty$ conforme $x \to 0$; algo que puede constituir una seria limitación con datos individuales.

Índices basados en la función de bienestar social

Los juicios de valor sobre la desigualdad siempre tienen algún contenido normativo, ya sea explícito o implícito. Esta es la razón por la que algunos autores prefieren partir de la interpretación de la desigualdad como una pérdida potencial en el bienestar colectivo y a partir de una **función de bienestar social**, que refleje de forma explícita los juicios de valor acerca de la relación entre desigualdad y bienestar, derivar una serie de índices de desigualdad que satisfagan ciertas propiedades; de hecho la mayoría de estadísticos de dispersión tradicionales pueden ser interpretados en términos de la función de bienestar social (Blackorby y Donalson (1978)). Es importante señalar que no existe ninguna indicación en esta literatura de que una igualdad completa es potencialmente alcanzable o incluso deseable, en general cualquier intento de conseguir una distribución más igualitaria reducirá el volumen total de renta disponible para repartir, por otra parte una igualdad completa es en ocasiones simplemente inalcanzable. Si bien fue Dalton (1920) el primer autor que argumentó que cualquier medida de desigualdad debe estar referida al bienestar económico, es la **familia de índices normativos de Atkinson** (1970) la más comúnmente utilizada (Kolm (1969, 1976a,b)).

Aunque no entraremos en la problemática teórica sobre la medición de la desigualdad y su relación con el bienestar conviene introducir algunos conceptos.[20] Dado un volumen de renta total, *Y*, Atkinson (1970) define la **renta igualitaria equivalente**, μ_e, como aquel nivel de renta *per capita* tal que, si fuese disfrutado por toda la población, generaría el mismo nivel

[18] Obsérvese que el logaritmo de la media geométrica es igual a la media aritmética de los logaritmos.

[19] Para valores no-negativos la media geométrica no es nunca mayor que la media aritmética, $\mu \geq \tilde{\mu}$, por tanto $T(0) \geq 0$. Cuando existe algún valor nulo la media geométrica es cero, aunque en este caso carece de utilidad como medida de posición, y para valores negativos no está definida (Kendall y Stuart (1963)).

[20] Deaton y Muellbauer (1980) Cap.9.

de bienestar social que la distribución inicial de renta. A partir de esta idea Atkinson (1970) define una medida de la desigualdad como

$$A = 1 - \frac{\mu_e}{\mu} \tag{9}$$

donde si la función de bienestar social es cóncava $\mu_e \leq \mu$; por tanto (9) recoge la pérdida de bienestar social generada por la desigual distribución de la renta.[21] En otras palabras, la medida de desigualdad de Atkinson (1970) no es más que el porcentaje de renta desperdiciada por la desigualdad existente y valorada en términos de una función de bienestar social, así por ejemplo, si $A = 0.3$ el índice de desigualdad de Atkinson nos indica que si la renta estuviera distribuida de forma igualitaria sólo necesitaríamos el 70% del volumen total del renta para alcanzar el mismo nivel de bienestar social.

Dar contenido operativo a (9) requiere postular una función de bienestar social. Atkinson (1970) restringe su análisis a la familia de funciones de bienestar social utilitaristas cuya función de valoración de la renta *per capita* tenga elasticidad constante. Esta familia de funciones viene dada por

$$W(x) = \begin{cases} \dfrac{1}{1-\varepsilon} \sum_i p_i x_i^{1-\varepsilon} & \varepsilon \neq 1 \\ \sum_i p_i \log x_i & \varepsilon = 1 \end{cases} \tag{10}$$

para $\varepsilon \geq 0$, lo que asegura la concavidad de $W(x)$. Dicha concavidad es estricta si $\varepsilon > 0$. Concavidad implica una preferencia por una distribución más igualitaria, por lo tanto $\varepsilon > 0$ indica aversión por la desigualdad.[22]

El **parámetro** ε se interpreta como el **grado de aversión relativa a la desigualdad** por parte de la sociedad y se supone **constante**. Para $\varepsilon = 0$ no existe ninguna aversión social a la desigualdad, la función de bienestar social es lineal y fue propuesta originariamente por Bentham (1907), a medida que aumenta ε aumenta el grado de aversión social a la desigualdad,

[21] La medida de desigualdad (9) supone, implícitamente, que las diferentes distribuciones de la renta no afectan a los incentivos individuales para alterar sus esfuerzos en orden a generar renta adicional, esto es, no afectan a la oferta de trabajo.

[22] La segunda igualdad en (10) se obtiene por aplicación de la regla de L'Hopital. Resulta interesante observar como en este caso, $\varepsilon = 1$, la función de bienestar social es igual al logaritmo de la media geométrica de la distribución de la renta *per capita*, $W(x) = \log \prod_i x_i^{p_i} = \log \tilde{\mu}$.

en el límite, cuando $\varepsilon \to \infty$, (10) tiende a la función de bienestar social propuesta por Rawls (1971) que valora el bienestar solamente del individuo más pobre.

A partir del concepto de **renta igualitaria equivalente** y la función de bienestar social (10) obtenemos la **familia de índices normativos de Atkinson** (1970), que viene dada por

$$A(\varepsilon) = \begin{cases} 1 - \left[\sum_i p_i \left(\dfrac{x_i}{\mu} \right)^{1-\varepsilon} \right]^{\frac{1}{1-\varepsilon}} & \varepsilon \neq 1 \\ 1 - \exp\left[\sum_i p_i \log\left(\dfrac{x_i}{\mu} \right) \right] = 1 - \prod_i \left(\dfrac{x_i}{\mu} \right)^{p_i} & \varepsilon = 1 \end{cases} \qquad (11)$$

para $\varepsilon \geq 0$. Obsérvese además que $A(1) = 1 - (\tilde{\mu}/\mu)$. Cuando $\varepsilon = 0$, es decir no hay aversión social a la desigualdad, $A(0) = 0$, por lo que el valor social de la desigualdad es nula cualquiera que sea la distribución de la renta. Para $\varepsilon > 0$ y distribuciones continuas el índice $A(\varepsilon)$ varía entre 0, igualdad perfecta, y 1, máxima desigualdad.

Puede demostrarse que $\dfrac{\partial A(\varepsilon)}{\partial \varepsilon} \geq 0$ de forma que $A(\varepsilon)$ nunca decrece al aumentar la aversión relativa a la desigualdad (Cowell (1995)).[23]

Una propiedad interesante de la **familia de índices normativos de Atkinson** (1970) proviene de la siguiente observación. Para $\varepsilon > 0$ y $\varepsilon \neq 1$

$$\sum_i p_i \left(\frac{x_i}{\mu} \right)^{1-\varepsilon} = \left[1 - A(\varepsilon) \right]^{1-\varepsilon} \qquad (12)$$

por lo que a partir de (3) y tomando $\beta = 1 - \varepsilon$ podemos escribir la **familia de índices de Theil** como

$$T(1-\varepsilon) = \frac{1}{\varepsilon(1-\varepsilon)} \left\{ 1 - \left[1 - A(\varepsilon) \right]^{1-\varepsilon} \right\} \qquad (13)$$

[23] No es posible obtener un resultado tan nítido para $T(\beta)$ en términos de β.

con lo que $T(1-\varepsilon)$ es una transformación monótonamente creciente del índice de Atkinson (1970), $A(\varepsilon)$, para un valor dado de $\varepsilon > 0$ y $\varepsilon \neq 1$. En consecuencia $A(\varepsilon)$ representa exactamente las mismas ordenaciones sobre diferentes vectores de renta, **x**, que la clase de medidas generalizadas de entropía, para un valor dado de ε y su correspondiente β.[24]

Para $\varepsilon = 1$ obtenemos el mismo resultado ya que (11) puede escribirse a partir de (6) como

$$A(1) = 1 - \exp(-T(0)) \tag{14}$$

por lo tanto

$$T(0) = -\log(1 - A(1)) \tag{15}$$

lo que muestra que $T(0)$ es una transformación monótonamente creciente de $A(1)$.

Por tanto $A(\varepsilon)$ y $T(\beta)$ para $\varepsilon > 0$ y $\beta = 1-\varepsilon$ son índices **ordinalmente equivalentes**, el *ranking* proporcionado por ambos índices para cualquier conjunto de curvas de Lorenz será el mismo. Sin embargo, dado que la relación entre ambos índices no es lineal no son cardinalmente equivalentes y, en consecuencia, no proporcionan la misma visión acerca de la reducción o amplificación de las desigualdades.

Otros índices

Aunque no realizaremos una discusión pormenorizada de estadísticos descriptivos que ayuden a caracterizar la distribución de una variable conviene introducir aquí algunas medidas comunes de dispersión que han sido utilizadas en el contexto de la medición de la desigualdad. Dos de las más habituales de estas medidas son, (i) el **rango**, $R(x) = \max\{x_i\}_{i=1}^{n} - \min\{x_i\}_{i=1}^{n}$, y (ii) la **varianza**. El **rango relativo** se define como la diferencia entre los valores máximo y mínimo dividido por la renta *per capita* media,

$$R_\mu(x) = \frac{1}{\mu}\left(\max\{x_i\}_{i=1}^{n} - \min\{x_i\}_{i=1}^{n}\right) \tag{16}$$

[24] Obsérvese que valores de $\varepsilon > 0$ implican $\beta < 1$, por tanto valores elevados de aversión a la desigualdad se corresponden con valores de β muy negativos, en particular $\varepsilon \to \infty$ es equivalente a $\beta \to -\infty$.

y toma valores en el intervalo $[0,1/p_i]$. Puesto que ignora todo lo que sucede entre los valores extremos, el rango no verifica el principio de las transferencias de Pigou-Dalton. Aunque (16) ha sido normalizado por μ otras normalizaciones son posibles, por ejemplo $R_{\min}(x)=\frac{\max\{x_i\}_{i=1}^n}{\min\{x_i\}_{i=1}^n}-1$, que toma valores en el intervalo $[0,\infty]$, o $R_{\max}(x)=1-\frac{\min\{x_i\}_{i=1}^n}{\max\{x_i\}_{i=1}^n}$, que toma valores en el intervalo $[0,1]$. Al parecer el interés por acotar los valores extremos en la distribución de la renta y la riqueza se remonta al menos hasta Platón (Saunders (1970)).

La **varianza**,

$$Var_\omega(x)=\sum_i p_i(x_i-\mu)^2 \tag{17}$$

si cumple el principio de las transferencias de Pigou-Dalton, al igual que la propiedad de independencia del tamaño de la población, pero viola la independencia respecto a la escala. Una forma de solucionar este problema es dividir la varianza por el cuadrado de la renta *per capita* media, lo que genera el **cuadrado del coeficiente de variación**,

$$CV_\omega(x)^2=\frac{Var_\omega(x)}{\mu^2} \tag{18}$$

que satisface las tres propiedades básicas mencionadas anteriormente. Obsérvese que para $\beta=2,\ T(2)=\frac{1}{2}CV_\omega(x)^2$, por lo que estos índices son **cardinalmente equivalentes**.

A diferencia de lo que sucede con G el coeficiente de variación, $CV_\omega(x)$, valora de forma uniforme las transferencias de renta dentro de la distribución, es decir la magnitud del cambio en $CV_\omega(x)$ ante transferencias de ricos a pobres es independiente de en que parte de la distribución se realizan dichas transferencias. Por ello esta no es una medida de desigualdad apropiada si queremos atribuir más peso al extremo inferior de la distribución, pero es un índice de dispersión perfectamente válido si no nos preocupa en que parte de la distribución se está produciendo la igualación de rentas.

Al contrario de lo que sucede con $Var_\omega(x)$, la **varianza de los logaritmos**[25] constituye un índice de desigualdad que es independiente de la escala,

[25] Todos los logaritmos son neperianos, es decir de base *e*.

$$Var_{\omega}(\log x) = \sum_i p_i (\log x_i - \log \tilde{\mu})^2 \qquad (19)$$

donde $\log \tilde{\mu} = \sum_i p_i \log x_i$. La varianza de los logaritmos tambien satisface la propiedad de independencia respecto al tamaño de la población. Sin embargo, no verifica el principio de las transferencias de Pigou-Dalton para la totalidad del domino de definición de rentas, en concreto dicho principio no se satisface para rentas superiores a $\tilde{\mu}e$, donde e es la base de los logaritmos neperianos (Cowell (1995)). Frente a esta característica indeseable, una propiedad interesante de $Var_{\omega}(\log x)$, como la varianza de cualquier variable, es que es descomponible en la suma de dos componentes, un componente *inter*-grupos y otro componente *intra*-grupos. Las propiedades de dicha descomposición serán brevemente analizadas más adelante. Otras características atractivas derivan de su relación con la distribución lognormal (Aitchison y Brown (1957)).

La **desviación de los logaritmos** de la renta *per capita* se realiza en ocasiones **respecto al logaritmo de la media aritmética**, $\log \mu$, en lugar de respecto al logaritmo de la media geométrica, $\log \tilde{\mu}$ (Sen (1973)). Esto genera una medida de desigualdad ligeramente diferente, la **varianza logarítmica**,

$$\nu_{\omega}(\log x) = \sum_i p_i (\log x_i - \log \mu)^2 \qquad (20)$$

que también es independiente de la escala. Al igual que sucede con $Var_{\omega}(\log x)$, $\nu_{\omega}(\log x)$ satisface la propiedad de independencia respecto al tamaño de la población pero tampoco verifica el principio de las transferencias de Pigou-Dalton para la totalidad del domino de definición de rentas, en concreto dicho principio no se satisface para rentas superiores a μe. (Cowell (1995)).

Puede comprobarse que

$$\nu_{\omega}(\log x) = Var(\log x) + (\log \mu - \log \tilde{\mu})^2 \qquad (21)$$

pero obsérvese que $(\log \mu - \log \tilde{\mu})$ es una medida de desigualdad, $T(0)$. Por lo tanto $\nu_{\omega}(\log x)$ es realmente la suma de dos medidas diferentes de desigualdad, $Var_{\omega}(\log x)$ y $T(0)^2$.

Tanto $Var_{\omega}(\log x)$ como $\nu_{\omega}(\log x)$ tienen una importante limitación, al igual que sucede con $T(0)$, no están definidas para distribuciones con renta cero.

En líneas generales podemos observar como las medidas de desigualdad propuestas en la literatura pueden entenderse como una media aritmética de funciones de distancia entre puntos, donde la ponderación es la frecuencia relativa. Dicho de otra forma, se trata de índices que son lineales en las frecuencias y convexos en las distancias.[26] Cuando la distancia se expresa como una función al cuadrado pero se desea mantener la comparabilidad con el nivel de la serie original se suele utilizar la raíz cuadrada del índice en cuestión. Como se mencionó en la introducción obsérvese que todos los índices son estadísticos ponderados, para obtener los correspondientes estadísticos no ponderados basta simplemente con sustituir p_i por $1/n$, $\forall i$.

Finalmente señalar que la aproximación seguida en este epígrafe ha sido meramente descriptiva y no ha tomado una fundamentación de inferencia estadística. Medidas de variabilidad, para datos muestrales, de muchos de los índices considerados pueden encontrarse en Kendall y Stuart (1963) u obtenerse aproximaciones asintóticas mediante el método delta (Rao (1973)) o técnicas de *bootstraping* (Mills y Zandvakili (1997)). En términos de la teoría estadística para inferir si los cambios en una distribución o curva de Lorenz son o no significativos los últimos años han visto un importante crecimiento de las técnicas disponibles a pesar de la complejidad del tema, una aplicación interesante de estas técnicas puede verse en Bishop, Formby y Thistle (1992).[27]

[26] Por el contrario las medidas de polarización propuestas por Esteban y Ray (1993, 1994) y Esteban (1996) son lineales en las distancias y convexas en las frecuencias. La razón es que la idea de polarización trata de poner énfasis de forma prioritaria en la similitud o disparidad entre los tamaños de las frecuencias relativas de los distintos puntos, es por ello que el concepto de polarización es adecuado en contextos multimodales.

[27] Sobre consideraciones teóricas acerca de como inferir la dominancia de una distribución o curva de Lorenz sobre otra pueden consultarse Gail y Gastwirth (1978), Beach y Davidson (1983), Gastwirth y Gail (1985), Beach y Kaliski (1986), Howes (1993), Bishop, Chow y Formby (1994), Bishop, Chakraborti y Thistle (1994) y Davidson y Duclos (1997). Una aplicación para España de estas técnicas puede verse en del Río y Ruíz-Castillo (1996).

3. EXPLORANDO EL ORIGEN DE LA DESIGUALDAD.

Una de las posibles aplicaciones de los índices de desigualdad es la de proporcionarnos información acerca de cuales son las causas de la misma, así como su importancia relativa. Este epígrafe aborda esta cuestión desde diversas perspectivas.

Descomponer un índice aditivamente en una serie de factores equivale a determinar que parte de la desigualdad total observada es atribuida a cada uno de esos factores. Para ello requerimos que los índices utilizados sean de naturaleza cardinal, es decir que su valor nos proporcione información no sólo acerca de un "mayor o menor" grado de desigualdad, sino tambien de "cuanto mayor o cuanto menor" grado de desigualdad. A continuación consideramos la descomponibilidad de un índice en varios sentidos.

Descomposición de un índice de desigualdad por subgrupos de población.

Este tipo de descomponibilidad consiste en la subdivisión de la población en grupos homogéneos, exhaustivos y mutuamente excluyentes, para analizar que parte de la desigualdad total es atribuible a cada uno de estos grupos. De acuerdo con esta idea un **índice** agregado **de desigualdad** se dice que **es aditivamente descomponible** en sentido débil **si puede ser escrito como la suma de un componente *inter*-grupos y un componente *intra*-grupos**, donde **(i)** el **componente *inter*-grupos**, que mide la **desigualdad externa**, es el valor del índice de desigualdad cuando cada miembro del grupo recibe la renta *per capita* media de dicho grupo, y **(ii)** el **componente *intra*-grupos**, que mide la **desigualdad interna**, es una suma ponderada de los índices de desigualdad para cada uno de los grupos, donde los pesos en la suma dependen sólo de las proporciones de renta y/o población de dicho grupo. La desigualdad interna es, por tanto, una suma ponderada de la desigualdad dentro de cada uno de los grupos, donde la ponderación trata de reflejar el peso de cada grupo dentro del total.

El índice de Gini, G, no es, en general, descomponible en el sentido anterior, sin embargo la familia de índices de Theil siempre es descomponible de acuerdo con la anterior definición, así como también, aunque en un sentido ligeramente diferente, la varianza de los logarítmos dada por (19). Finalmente el índice de Atkinson (1970) no es descomponible en el sentido mencionado. Aunque nuestra atención prioritaria se centrará en la familia de índices de Theil comentaremos brevemente sobre algunos de los resultados disponibles para el índice de Gini, así como la mencionada descomposición referente a la varianza de los logaritmos.

Sea $i = 1,2,...,n$ el índice para las unidades económicas de partida y $g = 1,2,...,G$ el índice que denota las agrupaciones de dichas unidades económicas de partida. Ahora denominamos n_g al número de unidades económicas de partida pertenecientes a la agrupación g, por tanto $N_g = \sum_{i \in n_g} N_i$, $Y_g = \sum_{i \in n_g} Y_i$, p_g es la población relativa respecto al total de la agrupación g, $p_g = \frac{N_g}{N} = \sum_{i \in n_g} p_i$, y_g es la renta relativa respecto al total de la agrupación g, $y_g = \frac{Y_g}{Y} = \sum_{i \in n_g} y_i$ y μ_g es la renta *per capita* media la agrupación g, $\mu_g = \frac{Y_g}{N_g} = \frac{1}{p_g} \sum_{i \in n_g} p_i x_i$.

De acuerdo con esta nomenclatura y la definición de descomponibilidad dada anteriormente un índice de desigualdad genérico, digamos I, es aditivamente descomponible en sentido débil si puede ser escrito como[28]

$$I = \sum_{g=1}^{G} \omega_g I_g + I_0 \tag{22}$$

donde $\omega_g = \omega_g(p_g, y_g)$, $g = 1,...,G$, son las ponderaciones de los índices de desigualdad dentro de cada uno de los grupos, I_g, que se utilizan para obtener el componente *intra*-grupos (desigualdad interna), $\Sigma_{g=1}^{G} \omega_g I_g$, e I_0 es el índice de desigualdad *inter*-grupos (desigualdad externa). A partir de esta descomposición la contribución *intra*-grupos (interna) es definida como el cociente entre el componente *intra*-grupos y el índice global, $\Sigma_{g=1}^{G} \omega_g I_g / I$; y la contribución *inter*-grupos (externa) es definida como el cociente entre el componente *inter*-grupos y el índice global, I_0 / I. El **componente *intra*-grupos (desigualdad interna),** $\Sigma_{g=1}^{G} \omega_g I_g$, nos indica el grado de desigualdad atribuible a las desigualdades dentro de las diversas agrupaciones, cuando no existen desigualdades entre las agrupaciones. Por su parte el **componente *inter*-grupos (desigualdad externa),** I_0, nos indica el grado de desigualdad que podemos atribuir a las desigualdades entre agrupaciones, independientemente de las desigualdades dentro de cada agrupación. Más adelante volveremos de nuevo sobre esta cuestión.

Puede demostrarse (Zagier (1983)) que el índice de Gini, G, para el total de la población satisface la siguiente desigualdad

[28] Obsérvese que esta descomposición puede anidarse si las agrupaciones se subdividen a su vez en grupos homogéneos, exhaustivos y mutuamente excluyentes, en este caso aparecerán tres términos en (22), ya que cada I_g se descompondrá en dos términos adicionales.

$$G \geq \sum_{g=1}^{G} p_g y_g G_g + G_0 \qquad (23)$$

El **primer término** de la parte derecha de la desigualdad (23) representa efectivamente el **componente *intra*-grupos (desigualdad interna)**, $\Sigma_{g=1}^{G} p_g y_g G_g$ es una suma ponderada de los índices de Gini para cada uno de los grupos, donde las ponderaciones son el producto de la proporción de población, p_g, por la proporción de renta, y_g, de cada grupo sobre el total. Por lo tanto las ponderaciones del componente *intra*-grupos, $\omega_g = p_g y_g$, suman menos de la unidad; obsérvese que la definición de descomponibilidad dada anteriormente no requiere que las ponderaciones sumen la unidad, es decir el componente *intra*-grupos no tiene por que ser una media ponderada de los índices de cada uno de los grupos, aunque como veremos a continuación ello genera problemas de interpretación a la hora de establecer la contribución porcentual a la desigualdad global de cada una de las agrupaciones. El **segundo término** de la parte derecha de la desigualdad (23) representa, por definición, el **componente *inter*-grupos (desigualdad externa)**, G_0, que no es más que el índice de Gini aplicado a partir de los datos agregados a las g agrupaciones consideradas.

La **desigualdad (23)** se convierte en **igualdad sólo cuando las distribuciones de los diferentes grupos no se solapan**, únicamente en este caso el índice de Gini puede descomponerse en la suma de un componente *intra*-grupos y un componente *inter*-grupos. Un ejemplo obvio de esta situación ocurre cuando la población es dividida en dos agrupaciones de acuerdo con un umbral de pobreza, de forma que tenemos un grupo de pobres y otro de no-pobres, esta característica ha facilitado la utilización del índice de Gini en la construcción de indicadores de pobreza (Sen (1976), Blackorby y Donalson (1980), Shorrocks (1995), Chakravarty (1983, 1990, 1997)), que cobran especial relevancia con la utilización de datos microeconómicos. Otra circunstancia en la que se produce esta situación es cuando $G_g = 0, \forall g$, de forma que toda la desigualdad proviene de las diferencias entre agrupaciones, en este caso $G = G_0$. Lo contrario no es, sin embargo, cierto, cuando la desigualdad externa es nula, $G_0 = 0$ porque $\mu_g = \mu \ \forall g$, y toda la desigualdad proviene de las diferencias dentro de las agrupaciones entonces, $G \geq \Sigma_{g=1}^{G} p_g y_g G_g$, ya que en este caso no se cumple la propiedad de no solapamiento entre las distribuciones de los diferentes grupos.

Así pues el índice de Gini no es, en general, descomponible aditivamente en el sentido señalado al principio de este epígrafe. Además, y aún en aquellos casos particulares en los que se verifica la mencionada descomponibilidad, el componente *intra*-grupos no puede ser definido como una media, y no simplemente una suma, ponderada por las proporciones de renta o población de los índices de Gini de cada uno de los grupos, al contrario de lo que sucede con algunos índices de la familia de índices de Theil. Sin embargo, todavía es posible acotar un poco más la desigualdad (23) en términos de la relación entre el índice de Gini para el total de la población y una media ponderada por las proporciones de renta o población de los índices de Gini para cada uno de los grupos; en este sentido es posible demostrar que (Zagier (1983))

$$G \geq \sum_{g=1}^{G} p_g G_g \tag{24}$$

$$G \geq \sum_{g=1}^{G} y_g G_g \tag{25}$$

Por lo tanto el índice de Gini es siempre mayor o igual que una media ponderada de los índices de Gini de cada uno de los grupos, tanto si consideramos ponderaciones por proporciones de renta como ponderaciones por proporciones de población. Una igualdad en sentido estricto en (24) y (25) se obtiene sólo cuando las distribuciones de los diferentes grupos son idénticas;[29] en cualquier otro caso, incluso si $G_0 = 0$, esto es, desigualdad externa nula, el índice de Gini para el conjunto de la población será estrictamente mayor que una media ponderada de los índices de Gini de cada uno de los grupos, ya consideremos proporciones de renta o proporciones de población como ponderaciones.

Al contrario de lo que sucede con *G*, **la familia de índices de Theil** (1967) **siempre es descomponible aditivamente de acuerdo con la definición anterior** (Shorrocks (1984)). Dicha descomposición adopta la forma

$$T(\beta) = \sum_{g=1}^{G} \omega_g T_g(\beta) + T_0(\beta) \tag{26}$$

donde $\omega_g = p_g^{1-\beta} y_g^{\beta}$, $g = 1,...,G$, son las ponderaciones de los índices de desigualdad dentro de cada uno de los grupos y que se utilizan para obtener el componente *intra*-grupos, $\Sigma_{g=1}^{G} \omega_g T_g(\beta)$, que es una suma ponderada de los índices de Theil para cada uno de los grupos; y $T_0(\beta)$ es el componente *inter*-grupos, esto es el índice de Theil aplicado a partir de los datos agregados a las *g* agrupaciones consideradas.

Examinemos con un poco de detalle los **tres casos** posibles:

[29] De forma que el agregado es simplemente una réplica de la distribución de cada uno de los grupos.

(1) Para $\beta = 1$ obtenemos ponderaciones que son proporciones de renta de cada agrupación sobre el total, $\omega_g = y_g$; por tanto sean dos agrupaciones con la misma desigualdad según $T(1)$, pero con distinto porcentaje de renta, la contribución a la desigualdad global será mayor para la agrupación que tenga un porcentaje de renta mayor, esto es la más rica.

En este caso la descomposición para el índice de desigualdad global, $T(1)$, que viene dado por (4), es la siguiente. Observando que el porcentaje de renta de cada unidad económica de partida sobre su correspondiente agrupación puede escribirse como $\frac{Y_i}{Y_g} = \frac{p_i x_i}{p_g \mu_g}$, entonces el **índice interno de desigualdad para cada agrupación** vendrá dado por

$$T_g(1) = \sum_{i \in n_g} \left(\frac{p_i x_i}{p_g \mu_g} \right) \log \left(\frac{x_i}{\mu_g} \right) \tag{27}$$

que de una forma análoga a (5) puede escribirse como

$$T_g(1) = \sum_{i \in n_g} \left(\frac{y_i}{y_g} \right) \log \left(\frac{y_i / y_g}{p_i / p_g} \right) \tag{28}$$

puesto que $\frac{x_i}{\mu_g} = \frac{y_i / y_g}{p_i / p_g}$.

De igual forma, observando que el porcentaje de renta de cada agrupación sobre el total puede escribirse como $y_g = \frac{Y_g}{Y} = \frac{p_g \mu_g}{\mu}$, el **índice de desigualdad externa entre agrupaciones** vendrá dado por

$$T_0(1) = \sum_{g} \left(\frac{p_g \mu_g}{\mu} \right) \log \left(\frac{\mu_g}{\mu} \right) \tag{29}$$

que de una forma análoga a (5) puede escribirse como

$$T_0(1) = \sum_{g} y_g \log \left(\frac{y_g}{p_g} \right) \tag{30}$$

puesto que $\frac{\mu_g}{\mu} = \frac{y_g}{p_g}$.

Ponderando los índices de Theil internos de cada agrupación, **utilizando como ponderación el porcentaje de renta de cada agrupación dentro del total,** $y_g = \frac{p_g \mu_g}{\mu}$, y sumando el índice de Theil externo entre agrupaciones se obtiene la descomposición (26) para $\beta = 1$, $T(1)$, que viene dada por (Esteban (1994))

$$T(1) = \sum_{g=1}^{G} \frac{p_g \mu_g}{\mu} T_g(1) + T_0(1) \tag{31}$$

o de una forma análoga a (5)

$$T(1) = \sum_{g=1}^{G} y_g T_g(1) + T_0(1) \tag{32}$$

(2) Para $\beta = 0$ obtenemos ponderaciones que son proporciones de población de cada agrupación sobre el total, $\omega_g = p_g$; por tanto sean dos agrupaciones con la misma desigualdad según $T(0)$, pero con distinta población, la contribución a la desigualdad global será mayor para la agrupación que esté relativamente más poblada.

En este caso la descomposición para el índice de desigualdad global, $T(0)$, que viene dado por (6), es la siguiente. Observando que el porcentaje de población de cada unidad económica de partida sobre su correspondiente agrupación puede escribirse como $\frac{N_i}{N_g} = \frac{p_i}{p_g}$, el **índice interno de desigualdad para cada agrupación** vendrá dado por

$$T_g(0) = -\sum_{i \in n_g} \left(\frac{p_i}{p_g} \right) \log \left(\frac{x_i}{\mu_g} \right) \tag{33}$$

De igual forma el **índice de desigualdad externa entre agrupaciones** vendrá dado por

$$T_0(0) = -\sum_{g} p_g \log \left(\frac{\mu_g}{\mu} \right) \tag{34}$$

Ponderando los índices de Theil internos de cada agrupación**, utilizando como ponderación el porcentaje de población de cada agrupación dentro del total**, p_g, y sumando el índice de Theil externo entre agrupaciones se obtiene la descomposición (26) para $\beta = 0$, $T(0)$, que viene dada por (Ruiz-Castillo (1986))

$$T(0) = \sum_{g=1}^{G} p_g T_g(0) + T_0(0) \tag{35}$$

(3) Para $\beta \neq 0{,}1$ obtenemos ponderaciones que son una combinación no lineal de proporciones de población y renta, $\omega_g = p_g^{1-\beta} y_g^{\beta}$. Si $\beta < 0$, la contribución de cada agrupación a la desigualdad global será tanto mayor cuanto más poblada esté en términos relativos y cuanto menor sea su porcentaje de renta sobre el total. Si $\beta > 1$, obtenemos el caso contrario, de forma que la contribución de cada agrupación a la desigualdad global será tanto mayor cuanto menos poblada esté en términos relativos y cuanto mayor sea su porcentaje de renta sobre el total. Para valores de $0 < \beta < 1$, la contribución de cada agrupación a la desigualdad global será tanto mayor cuando mayores sean p_g e y_g.

Obsérvese que cuando $\beta \neq 0{,}1$ los coeficientes $\omega_g = p_g^{1-\beta} y_g^{\beta}$ no suman la unidad, de forma que, al igual que sucedía con el índice de Gini (fórmula (23)), la contribución a la desigualdad del componente *intra*-grupos **no** es una media ponderada de la desigualdad dentro de cada uno de los grupos. Es posible demostrar que cuando $0 < \beta < 1$ entonces $\Sigma_{g=1}^{G} \omega_g$ es menor que la unidad, mientras que cuando $\beta < 0$ ó $\beta > 1$ entonces $\Sigma_{g=1}^{G} \omega_g$ es mayor que la unidad. Además de las dificultades obvias de interpretación a la hora de establecer la contribución porcentual a la desigualdad global de cada una de las agrupaciones, lo más problemático en este caso es que puede demostrarse (Theil (1967)) que $1 - \Sigma_{g=1}^{G} \omega_g$ es proporcional al término $T_0(\beta)$, lo que da lugar a posibles ambigüedades de interpretación al variar los coeficientes de ponderación, $\omega_g = p_g^{1-\beta} y_g^{\beta}$, si modificamos $T_0(\beta)$. Esto dificulta la interpretación de la descomposición en ciertos contextos y por esta razón la descomposición por subgrupos de población no parece adecuada para valores de $\beta \neq 0{,}1$.

No obstante la descomposición para el índice de desigualdad global, $T(\beta)$, que viene dado por (3), es la siguiente. El **índice interno de desigualdad para cada agrupación** vendrá dado por

$$T_g(\beta) = \frac{1}{\beta(\beta-1)} \sum_{i \in n_g} \left(\frac{p_i}{p_g}\right) \cdot \left[\left(\frac{x_i}{\mu_g}\right)^{\beta} - 1\right] \quad \beta \neq 0{,}1 \tag{36}$$

De igual forma, el **índice de desigualdad externa entre agrupaciones** vendrá dado por

$$T_0(\beta) = \frac{1}{\beta(\beta-1)} \sum_{g} p_g \left[\left(\frac{\mu_g}{\mu}\right)^{\beta} - 1\right] \quad \beta \neq 0{,}1 \tag{37}$$

Ponderando los índices de Theil internos de cada agrupación**, utilizando como ponderación** $\omega_g = p_g^{1-\beta} y_g^{\beta}$, y sumando el índice de Theil externo entre agrupaciones se obtiene la descomposición (26) para $\beta \neq 0{,}1$, $T(\beta)$, que viene dada por (Shorrocks (1980, 1984))

$$T(\beta) = \sum_{g=1}^{G} p_g^{1-\beta} y_g^{\beta} T_g(\beta) + T_0(\beta) \tag{26}$$

y de la cual los casos para $\beta = 1$ ó $\beta = 0$ son casos particulares.

Prometimos anteriormente volver sobre las cuestiones de **interpretación** ya que puesto que las medidas de desigualdad descomponibles se utilizan para cuantificar la contribución de cada factor a la desigualdad global éstas son de vital importancia. Al igual que en (22) la interpretación de los dos términos en la descomposición (26) es la siguiente. El **componente *intra*-grupos (desigualdad interna),** $\Sigma_{g=1}^{G} \omega_g T_g(\beta)$, nos indica el grado de desigualdad atribuible a las desigualdades dentro de las diversas agrupaciones, cuando no existen desigualdades entre las agrupaciones. Por su parte el **componente *inter*-grupos (desigualdad externa),** $T_0(\beta)$, nos indica el grado de desigualdad que podemos atribuir a las desigualdades entre agrupaciones, independientemente de las desigualdades dentro de cada agrupación. Esta interpretación puede comprenderse mejor con un **ejemplo**, lo que nos ayudará a descubrir posibles inconsistencias. Consideremos a las provincias como unidades económicas de partida que se agrupan de forma natural en unidades geográficas político-administrativas más amplias, las comunidades autónomas, por ello puede resultar de interés preguntarse hasta que punto la desigualdad observada en la distribución provincial de la renta es debida a desigualdades entre CCAA o a desigualdades *inter*-provinciales dentro de cada CCAA. Si la desigualdad agregada se explica fundamentalmente por las diferencias entre CCAA, entonces sería recomendable diseñar de forma centralizada políticas de desarrollo regional, así como un mecanismo de transferencias entre CCAA que tendiera a reducir las desigualdades. Por el contrario, si la desigualdad agregada se explica fundamentalmente por las desigualdades *inter*-provinciales internas a cada CCAA, la necesidad de políticas de redistribución específicas dentro de cada CCAA cobrarían mayor importancia, en este caso tambien resulta de interés el examen individualizado de los índices de desigualdad de cada CCAA, ya que estas pueden no presentar un comportamiento homogéneo.

La pregunta concreta de **cuanta desigualdad puede atribuirse a las diferencias entre CCAA en España** puede interpretarse, como argumenta Shorrocks (1980), de dos formas diferentes:[30]

(i) **¿Cual sería el nivel de desigualdad global observado si las diferencias entre CCAA (agrupaciones) fueran las únicas que existieran?.** La respuesta a esta pregunta sugiere comparar el nivel de desigualdad global con el que se observaría si la desigualdad **dentro** de cada CCAA (agrupación) fuera nula, es decir desigualdad interna nula, pero se mantuvieran las diferencias en renta *per capita* entre CCAA (agrupaciones). Para índices aditivamente descomponibles en el sentido antes mencionado esto eliminaría el componente *intra*-grupos (desigualdad interna), ya que $T_g(\beta) = 0 \; \forall g$, y dejaría sólo el componente *inter*-grupos (desigualdad externa). Por lo tanto, en este caso, para cualquier valor de β, la respuesta vendría dada de forma inequívoca por el término $T_0(\beta)$ en (26), ya que las redistribuciones dentro de cada agrupación no afectan a la media de dicha agrupación y en consecuencia tampoco al componente *inter*-grupos.

(ii) Alternativamente podríamos preguntarnos: **¿En cuanto se reduciría la desigualdad global si se eliminasen las diferencias entre CCAA (agrupaciones), manteniéndose la desigualdad existente dentro de cada una de ellas?.** La respuesta a esta pregunta sugiere comparar el nivel de desigualdad global con el que se observaría si igualaramos la renta *per capita* de todas las CCAA (agrupaciones), es decir desigualdad externa nula, pero mantuviéramos intacta la desigualdad existente dentro de cada CCAA (agrupación). Para índices aditivamente descomponibles en el sentido antes mencionado esto eliminaría el componente *inter*-grupos (desigualdad externa); con lo que aparentemente la respuesta vendría dada de nuevo por $T_0(\beta)$. Sin embargo esta intuición no es totalmente correcta, la razón es que el resultado de este experimento conceptual no es invariante a la elección de β, es decir la reducción en la desigualdad no viene dada por $T_0(\beta)$ para cualquier β. Ello se debe a que los coeficientes de ponderación de la desigualdad experimentada dentro

[30] Sobre cuestiones de interpretación ver también Davies y Shorrocks (1978).

de cada agrupación, $\omega_g = p_g^{1-\beta} y_g^{\beta}$, pueden no ser independientes del procedimiento utilizado para eliminar la desigualdad entre agrupaciones. En otras palabras, ω_g puede no ser independiente del procedimiento para hacer nulo el término $T_0(\beta)$ en (26), ya que la redistribución entre grupos afectará, en general, a ω_g y por tanto a la contribución del componente *intra*-grupos.

En concreto, para el caso de nuestro ejemplo, igualar la renta *per capita* de todas las CCAA implica un mecanismo de transferencias entre las CCAA, este cambio en la renta de cada agrupación, que se distribuye entre las unidades económicas de partida de forma que $\mathrm{T}_g(\beta)$ en (26) no se altere, afectará tambien a la proporción de renta de cada agrupación sobre el total, y_g, y, por consiguiente, a $\omega_g = p_g^{1-\beta} y_g^{\beta}$ a menos que este coeficiente sea independiente de y_g, lo que sólo sucede cuando $\beta = 0$; en este caso los coeficientes de ponderación ω_g vienen dados por la importancia demográfica relativa de las diferentes agrupaciones, $\omega_g = p_g$. Por tanto, **sólo para $\beta = 0$ la respuesta a ambas preguntas coincide y es igual al término $T_0(\beta)$ en (26)**, o lo que es lo mismo a $T_0(0)$ en (35).

Este argumento apoya, en este caso particular, la utilización de ponderaciones por las proporciones de población y no por las proporciones de renta en la descomposición por subgrupos de población. Obsérvese que el argumento anterior en favor de la utilización de los índices para $\beta = 0$, ponderaciones según la importancia demográfica relativa, descansa en buscar una descomposición tal que las ponderaciones utilizadas para obtener el componente *intra*-grupos, $\omega_g = \omega_g(p_g, y_g)$, sean independientes del procedimiento para eliminar la desigualdad entre agrupaciones, es decir del procedimiento para hacer nulo el término $T_0(\beta)$ en (26). En nuestro ejemplo hemos supuesto que dicho procedimiento requiere un mecanismo de transferencias de renta y son estas transferencias las que alteran $\omega_g = \omega_g(p_g, y_g)$, a menos que dichas ponderaciones sean independientes de y_g. Si hubieramos supuesto que es factible realizar transferencias de población para hacer nulo el término $T_0(\beta)$ en (26), entonces el razonamiento anterior hubiera apoyado la utilización de los índices para $\beta = 1$, ponderaciones por proporciones de renta, ya que en esta situación este tipo de transferencias alteran $\omega_g = \omega_g(p_g, y_g)$, a menos que dichas ponderaciones sean independientes de p_g. Por lo tanto es la estructura del problema la que determina cual es el valor de β que nos proporciona una interpretación adecuada en cada caso.

Finalmente dado el resultado estadístico de que la **varianza** siempre es **descomponible en un componente *intra*-grupos y un componente *inter*-grupos**, consideraremos la descomposición por subgrupos de población de la **varianza de los logaritmos** de x_i, por ser este índice independiente respecto a la escala de medida. A partir de (19) podemos escribir

$$\begin{aligned} Var_{\omega}(\log x) &= \sum_{g}\sum_{i\in n_g} p_i(\log x_i - \log\tilde{\mu})^2 \\ &= \sum_{g}\sum_{i\in n_g} p_i\left[(\log x_i - \log\tilde{\mu}_g) + (\log\tilde{\mu}_g - \log\tilde{\mu})\right]^2 \end{aligned} \tag{38}$$

donde $\tilde{\mu}_g$ es la media geométrica de la agrupación g, $\tilde{\mu}_g = \prod_{i\in n_g} x_i^{(p_i/p_g)}$.

Expandiendo el cuadrado de la parte derecha de la igualdad en (38) y operando en los tres términos resultantes obtenemos que (19) puede escribirse como la suma de dos componentes

$$Var_{\omega}(\log x) = \sum_{g} p_g Var_{g\omega}(\log x) + Var_{0\omega}(\log\tilde{\mu}_g) \tag{39}$$

el primer término de la parte derecha de la igualdad es el **componente *intra*-grupos,** una **media ponderada por las proporciones de población**, p_g, **de las varianzas internas de las diferentes agrupaciones**, $Var_{g\omega}(\log x) = \sum_{i\in n_g}\left(\frac{p_i}{p_g}\right)(\log x_i - \log\tilde{\mu}_g)^2$, y el segundo término es el **componente *inter*-grupos**, la **varianza externa entre agrupaciones**, $Var_{0\omega}(\log\tilde{\mu}_g) = \sum_{g} p_g(\log\tilde{\mu}_g - \log\tilde{\mu})^2$.

La descomposición de la varianza de los logaritmos utiliza pues ponderaciones según la estructura demográfica y no está sujeta a los problemas mencionados anteriormente, de forma que las dos interpretaciones posibles del componente *inter*-grupos coinciden en este caso. Una advertencia es, sin embargo, necesaria; ahora, para calcular el término $Var_{0\omega}(\log\tilde{\mu}_g)$ se elimina la desigualdad dentro de cada agrupación asignando a cada unidad económica de partida la media geométrica de la agrupación a la que pertenece, así pues los valores medios de la distribución van referidos a medias geométricas, no aritméticas como en el caso de los índices de Theil. Esta peculiar forma de eliminar la desigualdad dentro de cada agrupación hace que la descomposición de la varianza de los logaritmos **no** sea aditivamente descomponible en el sentido definido al principio de este epígrafe, (22). La razón estriba en que igualar la renta *per capita* de cada miembro de la agrupación g, digamos, mediante un sistema de transferencias, manteniendo μ_g constante afectará tanto a $\tilde{\mu}_g$ como a $\tilde{\mu}$, y por tanto tambien a sus logaritmos, en consecuencia el término *inter*-grupos, $Var_{0\omega}(\log\tilde{\mu}_g)$, no es independiente de la distribución dentro de cada agrupación.

La descomposición de la varianza (39) se verifica para cualquier variable, ya sea el logaritmo de la renta *per capita*, log x_i, o el valor original, x_i. La razón para elegir el logaritmo de x_i era considerar un índice que fuera independiente respecto a la escala. Como ya observamos anteriormente una transformación alternativa que permite obtener un índice con esta propiedad es x_i/μ. La varianza de x_i/μ es el **cuadrado del coeficiente de variación**, (18), cuya descomposición puede ser obtenida a partir de (39).

$$CV_{\omega}(x)^2 = \frac{Var_{\omega}(x)}{\mu^2} = \frac{\sum_g p_g Var_{g\omega}(x) + Var_{0\omega}(\mu_g)}{\mu^2} \tag{40}$$

siendo μ_g la media aritmética de la agrupación *g*.

$$\begin{aligned} CV_{\omega}(x)^2 &= \sum_g \frac{p_g \mu_g^2}{\mu^2} . \frac{Var_{g\omega}(x)}{\mu_g^2} + CV_{0\omega}(\mu_g)^2 \\ &= \sum_g \left(\frac{y_g^2}{p_g} \right) CV_{g\omega}(x)^2 + CV_{0\omega}(\mu_g)^2 \end{aligned} \tag{41}$$

puesto que $\mu_g = \frac{Y_g}{N_g} = \frac{y_g}{p_g}\mu$.

En (41) el **componente *inter*-grupos** es el **cuadrado del coeficiente de variación externo referido a las medias aritméticas de cada agrupación** y el **componente *intra*-grupos** es una **suma ponderada de los coeficientes de variación internos al cuadrado de las diferentes agrupaciones**, sin embargo, las ponderaciones de esta suma, que dependen de las proporciones de renta y población, no suman la unidad.[31]

En **resumen**, si bien en general el índice de Gini y la varianza de los logaritmos no son aditivamente descomponibles en el sentido definido en este epígrafe los índices de Theil (1967) para $\beta = 0$ ó $\beta = 1$ se descomponen aditivamente de forma adecuada en la suma de dos componentes, un componente *inter*-grupos y un componente *intra*-grupos, siendo este último una media ponderada de los índices de desigualdad para cada una de las agrupaciones. Sin embargo las contribuciones relativas de cada uno de los componentes será, en general, diferente para cada uno de los estadísticos mencionados. Cada índice enfatiza diferentes

[31] Este resultado podía haberse obtenido directamente a partir de la observación $T(2) = \frac{1}{2} CV_{\omega}(x)^2$.

aspectos de la distribución de x_i y, por tanto, no hay razón para esperar que la contribución relativa de ambos componentes en todos los índices sea la misma.

Por ejemplo, puesto que $T(0)$ no está definido para distribuciones con renta cero, $T(0) \to +\infty$ conforme $x_i \to 0$ para algún i, la descomposición para este índice será muy sensible a cambios en los niveles bajos de renta, mientras que esto no sucederá con $T(1)$. Así, el componente *intra*-grupos (interno) medido mediante la descomposición de $T(0)$, (35), puede hacerse arbitrariamente grande en relación al componente *inter*-grupos (externo) redistribuyendo renta dentro de una agrupación de forma que la renta *per capita* de una unidad económica de partida de esta agrupación tienda a cero.[32] Por tanto la contribución *inter*-grupos puede hacerse tan pequeña como se desee simplemente haciendo que la renta de una unidad económica de partida tienda a cero. El mismo comentario es aplicable a la descomposición para $Var_\omega(\log x)$, aunque este índice no verifique la definición de descomponibilidad utilizada en este epígrafe.

Las descomposiciones analizadas son puramente descriptivas y no están basadas en consideraciones estadísticas, de hecho la familia de índices de Theil no posee una distribución estadística conocida. No obstante, en principio la descomposición de la varianza podría fundamentarse en el análisis de varianza estándar, de forma que además de cuantificar la contribución del componente *inter*-grupos sería tambien posible contrastar su significación estadística.

Desagregación de la renta en factores aditivos.

Consideraremos ahora la descomponibilidad de un índice de desigualdad según la fuente de renta de la economías domésticas, las regiones o los países. Así por ejemplo podemos considerar la descomposición de la renta entre rentas del trabajo y del capital, o la desagregación del VAB por sectores productivos, y tratar de descomponer la desigualdad global en la desigualdad atribuible a cada uno de los factores que componen la renta agregada. En concreto consideraremos que la renta puede considerarse como el agregado de K factores

$$Y_i = \sum_{k=1}^{K} Y_{ki} \tag{42}$$

[32] En esta situación una transferencia de una unidad económica de partida pobre a otra más rica, dentro de la misma agrupación, no afectará al componente *inter*-grupos, mientras que incrementará el componente *intra*-grupos y el valor global del índice.

con lo que podemos escribir la renta *per capita* global como

$$x_i = \frac{Y_i}{N_i} = \frac{\Sigma_{k=1}^{K} Y_{ki}}{N_i} = \Sigma_{k=1}^{K} x_{ki} \tag{43}$$

donde x_{ki} es la renta *per capita* asociada al factor *k*. Obsérvese que esta desagregación no altera las ponderaciones de cada factor respecto a las del agregado.

Determinar que parte de la desigualdad en la distribución de la renta *per capita*, $\mathbf{x} = (x_1, x_2, \ldots, x_n)'$, corresponde al factor *k* exige asignar a dicho factor no sólo sus efectos directos sobre la desigualdad global sino tambien sus efectos indirectos, estos últimos pueden ser positivos o negativos, y dado que no existe una forma única de asignar los efectos indirectos la descomposición de un índice particular, caso de ser posible, no será en general única. Por otra parte si la contribución del factor *k* a la desigualdad global es la suma de sus efectos directos e indirectos, la desigualdad global no será, en general, igual a la suma, simple o ponderada, de las desigualdades en cada uno de los factores. Estos problemas hacen que muchas medidas de desigualdad no permitan una descomposición nítida por factores aditivos de renta.

La dificultades asociadas a este tipo de descomposición pueden ilustrarse facilmente utilizando la **varianza** como índice de desigualdad,[33]

$$Var_{\omega}(x) = Var_{\omega}\left(\Sigma_{k=1}^{K} x_k\right) = \sum_k Var_{\omega}(x_k) + \sum_k \sum_{j \neq k} Cov_{\omega}(x_k, x_j) \tag{44}$$

Si los diferentes factores de renta estuviesen incorrelacionados (44) se convertiría en

$$Var_{\omega}(x) = \sum_k Var_{\omega}(x_k) \tag{45}$$

con lo que la contribución del factor *k* a la desigualdad global vendría dada por $Var_{\omega}(x_k)$. Es la presencia de correlación entre los diversos factores que componen la renta la que genera el problema en la descomposición ya que deberemos determinar como la interacción entre dichos factores, los términos $Cov_{\omega}(x_k, x_j)$ en (44), son distribuidos entre las contribuciones individuales de los distintos factores, y puesto que no existe una forma única de efectuar esta distribución no existirá una descomposición factorial única para $Var_{\omega}(x)$.

[33] El subíndice *k* hace referencia a los factores que componen la renta.

Aunque volveremos más adelante sobre la descomposición de $Var_\omega(x)$ examinaremos en primer lugar los resultados para el **índice de Gini**. En particular estaremos interesados en determinar la relación entre *G* para la renta *per capita* agregada y *G* para sus componentes. A partir de la definición de *G*, (1), podemos escribir

$$
\begin{aligned}
2\mu G(x) &= \sum_i \sum_j p_i p_j \left| \sum_k x_{ki} - \sum_k x_{kj} \right| \\
&= \sum_i \sum_j p_i p_j \left| \sum_k (x_{ki} - x_{kj}) \right| \leq \sum_k \sum_i \sum_j p_i p_j \left| x_{ki} - x_{kj} \right|
\end{aligned} \tag{46}
$$

utilizando la propiedad $|\Sigma_i a_i| \leq \Sigma_i |a_i|$. A partir de la definición del índice de Gini para cada uno de los factores de renta,

$$G(x_k) = \frac{1}{2\mu_k} \sum_i \sum_j p_i p_j \left| x_{ki} - x_{kj} \right| \tag{47}$$

donde $\mu_k = Y_k / N = \Sigma_i p_i x_{ki}$ y $Y_k = \Sigma_{i=1}^n Y_{ki}$; y de la observación de que la renta *per capita* media para el agregado puede obtenerse como suma de las rentas *per capita* medias para cada uno de los factores, $\mu = \Sigma_k \mu_k$, podemos escribir $2\mu G(x) \leq 2 \sum_k \mu_k G(x_k)$, es decir

$$G(x) \leq \sum_k \frac{\mu_k}{\mu} G(x_k) \tag{48}$$

Este resultado nos dice que **el índice de Gini para el agregado es menor o igual que una media ponderada de los índices de Gini para cada uno de los factores**, donde las ponderaciones son las proporciones medias de cada factor en el agregado, $\mu_k / \mu = Y_k / Y$. Por tanto la información sobre la distribución de los diferentes factores nos permite obtener una cota superior a la desigualdad global observada en la distribución de la renta *per capita* agregada, cuando dicha desigualdad es medida a través del índice de Gini.[34] Puede demostrarse que cuanto mayor sea la correlación entre los *rankings* de la renta *per capita* agregada y sus factores más cercano estará *G*(*x*) de $\sum_k \frac{\mu_k}{\mu} G(x_k)$ y que una igualdad en sentido estricto se

[34] Debe observarse que estamos interesados en descomposiciones factoriales no triviales, en el sentido de que la contribución a la desigualdad de un factor concreto no es independiente de como se distribuye dicho factor. Descomposiciones triviales, cuando la contribución del factor a la desigualdad global es independiente de su distribución, generan resultados poco interesantes, en concreto en este caso la contribución porcentual a la desigualdad global de cada factor será igual a la proporción de dicho factor en el agregado, ya que consideraciones distribucionales no son de importancia (Shorrocks (1982)).

obtiene **si y sólo si** el *ranking* para cada una de las agrupaciones o individuos, *i*, es el mismo para cada uno de los factores, *k*, que componen la renta.

Esta última observación nos permite obtener lo que Shorrocks (1982) llama la descomposición "natural" del índice de Gini y que fue inicialmente propuesta por Fei, Ranis y Kuo (1978). Ordenamos los factores de renta de acuerdo con el *ranking* proporcionado por la renta *per capita* agregada de las diversas agrupaciones o individuos, de forma que llamando $\mathbf{x}_k = (x_{k1}, x_{k2}, \dots, x_{kn})'$ a los vectores "ordenados" de los factores se verifica $\mathbf{x} = \Sigma_{k=1}^{K} \mathbf{x}_k$, en un sentido matricial. Obsérvese que la ordenación en $\mathbf{x}_k \ \forall k$ se corresponde con la derivada de **x**. En este caso podemos escribir

$$G(x) = \sum_k \frac{\mu_k}{\mu} \overline{G}(x_k) \tag{49}$$

donde $\overline{G}(x_k)$ es un **pseudo-índice de Gini** calculado para el factor *k* a partir de la ordenación proporcionada por el agregado, **x**. Sólo cuando dicha ordenación coincida con el *ranking* para el factor *k* entonces se verificará que $G(x_k) = \overline{G}(x_k)$ y cuando esto se cumpla para todos los factores la desigualdad (48) se verificará en forma de igualdad. La ecuación (49) proporciona una descomposición consistente para el índice de Gini si definimos la contribución a la desigualdad global del factor *k* como $\frac{\mu_k}{\mu} \overline{G}(x_k)$.

Debe observarse que aunque no aparecen términos de interacción en la descomposición (49) estos efectos han sido implícitamente asignados entre las contribuciones de los diferentes factores al ordenar, de forma un tanto artificial, los elementos de $\mathbf{x}_k$ de acuerdo con el *ranking* de **x**. El problema es que la asignación de dichos efectos no es explícita a partir de la obtención de (49), pero ello sugiere que pueden existir otras descomposiciones alternativas que puedan resultar más atractivas en determinadas situaciones.

Desde el punto de vista computacional es necesario calcular $\overline{G}(x_k)$ para que (49) sea operativa. Para datos de individuos u hogares el índice de Gini puede escribirse en función de la renta de cada indiviudo y su *ranking* asociado (Sen (1973), Fei, Ranis y Kuo (1978), Milanovic (1997)), por lo que en este caso $\overline{G}(x_k)$ podría obtenerse a partir de una fórmula concreta de cálculo (Fei, Ranis y Kuo (1978), Shorrocks (1982)); para datos procedentes de agrupaciones, como son los utilizados en nuestra aplicación, $\overline{G}(x_k)$ debe ser calculado por integración numérica del área bajo la **pseudo-curva de Lorenz** (Plotnick (1981)) asociada al pseudo-índice de Gini correspondiente. La construcción de esta pseudo-curva es sencilla y sigue los mismos principios que ya expusimos al examinar la curva de Lorenz en la sección anterior, teniendo en cuenta ahora que el vector representativo de factor *k*, $\mathbf{x}_k$, ha sido ordenado de acuerdo con el *ranking* proporcionado por **x**. Una vez realizada esta salvedad recordamos que $F_s = \Sigma_{i=1}^{s} p_i$ es la proporción de población que recibe una renta *per capita* igual o inferior a x_s, y definimos $\Phi_{ks} = \Sigma_{i=1}^{s} y_{ki}$, siendo $y_{ki} = Y_{ki} / Y_k$, como la proporción del factor *k* que reciben las agrupaciones con nivel de renta *per capita* igual o inferior a x_s. Definiendo $F_0 = \Phi_{k0} = 0$, entonces la relación entre Φ_{ks} y F_s es la llamada pseudo-curva de Lorenz[35] para una distribución dada de la renta *per capita* y de su factor *k*. Una representación gráfica, para distribuciones continuas, de la pseudo-curva y curva de Lorenz para el factor *k* es ofrecida en el **gráfico 3**. Es posible demostrar que la pseudo-curva de Lorenz no estará nunca por debajo de la curva de Lorenz del factor correspondiente (Plotnick (1981)),[36] y por tanto que el área entre entre la pseudo-curva de Lorenz y la linea de igualdad perfecta no será nunca mayor que el área entre la curva de Lorenz y la linea de igualdad perfecta.[37] Definimos ahora, desde el punto de vista geométrico, el **pseudo-índice de Gini** como **cociente del área entre la pseudo-curva de Lorenz y la linea de igualdad perfecta dividida por el área del triángulo 0AB**, que es igual a ½; por tanto $\overline{G}$ es equivalente a **dos veces el área comprendida entre la pseudo-curva de Lorenz y la linea de igualdad perfecta**. Alternativamente, el **pseudo-índice de Gini** es igual a **1 menos dos veces el área bajo la pseudo-curva de Lorenz**. La integración numérica de estas áreas proporciona un método adecuado para el cálculo de los pseudo-índices correspondientes, lo que permite hacer operativa la descomposición (49).

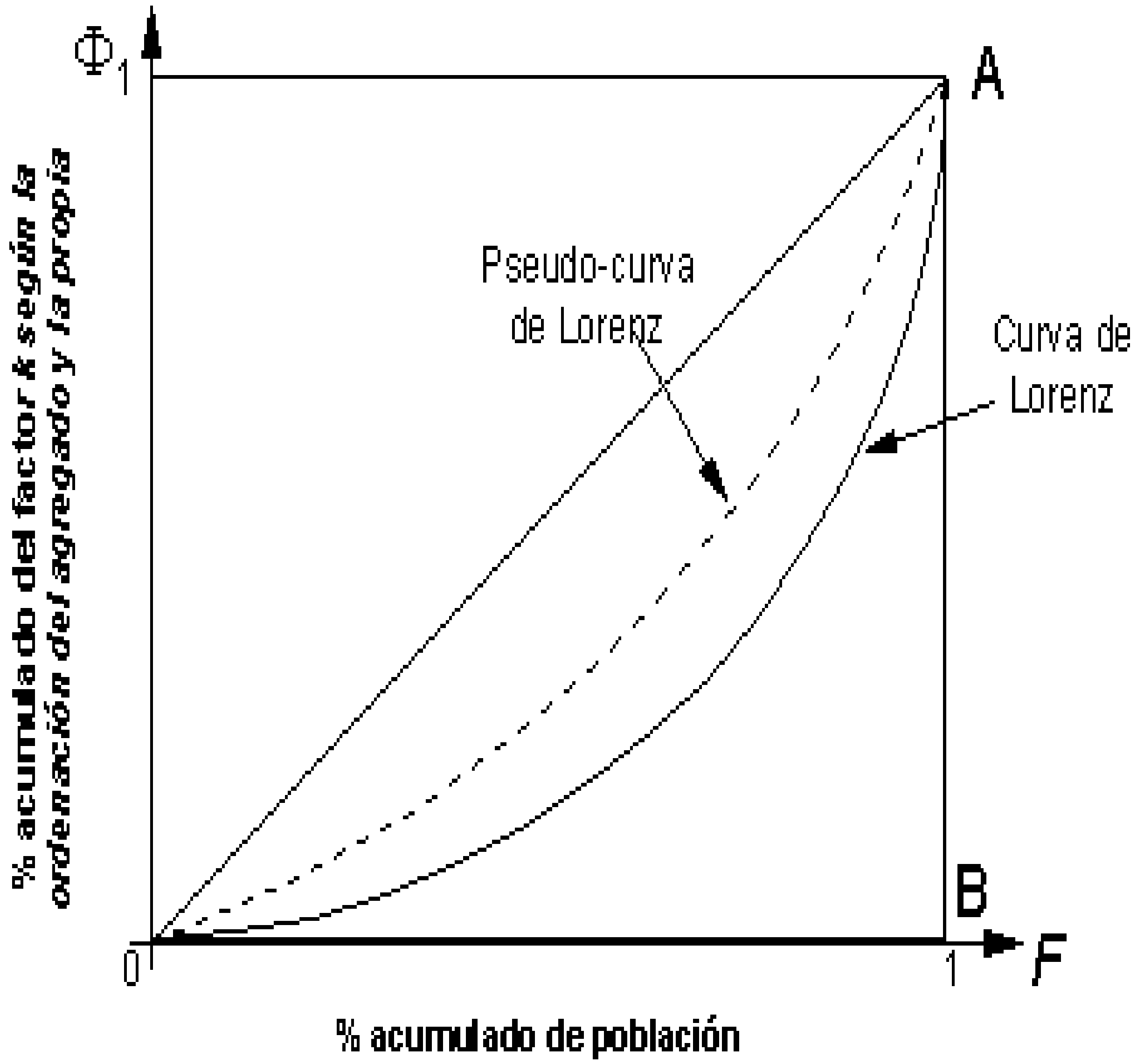
---

[35] Algunos autores utilizan la terminología **curvas de concentración** (Davidson y Duclos (1997)).

[36] Esta afirmación es cierta para distribuciones continuas o discretas en el caso de datos individuales, en esta situación los valores representados en el eje de abscisas son los mismos para ambas curvas, sin embargo no se cumple para datos de agrupaciones cuando los porcentajes de población de cada agrupación son diferentes y los puntos intermedios se obtienen por interpolación lineal, ya que en este caso los valores representados en el eje de abscisas son diferentes para cada una de las curvas. No obstante, aún en este caso, el área entre la pseudo-curva de Lorenz y la línea de igualdad perfecta no será nunca mayor que el área entre la curva de Lorenz y la línea de igualdad perfecta.

[37] Adicionalmente es posible demostrar que el área entre ambas curvas varía entre 0, cuando no hay alteración en los *rankings*, y el índice de Gini para el factor *k*, es decir dos veces el área comprendida entre la curva de Lorenz para dicho factor y la línea de igualdad perfecta, cuando se produce una inversión completa en los *rankings* de los individuos o agrupaciones.

Gráfico 3: Curva y pseudo-curva de Lorenz del factor *k*



La utilización de la descomposición "natural" de *G*, (49), para examinar la contribución de los diversos factores de renta a la desigualdad global requiere de dos condiciones para que tenga sentido. En primer lugar el índice de Gini debe considerarse una medida de desigualdad apropiada, lo cual es adecuado en vista de sus propiedades, y en segundo lugar, la regla de descomposición "natural", que procede de la forma en la que el índice de Gini es obtenido (geométricamente), debe ser aceptable; en otras palabras la ordenación de los diferentes factores de acuerdo con el *ranking* proporcionado por la renta *per capita* agregada debe ser razonable. Esta última condición depende del problema concreto que estemos tratando pero en general no es adecuada. Aplicaciones concretas donde puede ser útil mantener el *ranking* proporcionado por la renta *per capita* agregada pueden encontrarse en la literatura sobre los efectos de la imposición y la medición de la equidad de un sistema redistributivo (Kakwani (1977), Atkinson (1980), Plotnick (1981,1982), Duclos (1993), Lerman y Yitzhaki (1995), Rabadán y Salas (1996)), donde la renta *per capita* agregada debe entenderse antes de impuestos; sin embargo son difíciles de encontrar en otros contextos, en los que el mantenimiento de la ordenación para los diversos factores según el *ranking* proporcionado por la renta *per capita* agregada es en general una arbitrariedad sin posible justificación.

Resulta interesante ahora volver de nuevo sobre el problema de la descomposición de $Var_{\omega}(x)$ a partir de (44). Determinar la contribución de cada factor a la desigualdad global implica establecer una regla de asignación de los efectos de interacción entre los factores que componen la renta, los términos $Cov_{\omega}(x_k, x_j)$, a la contribución individual de cada factor. En ausencia de información adicional parece razonable asignar a cada factor, digamos *k*, la mitad del valor de todos los efectos interacción que implican a este factor, es decir la mitad de todos los términos $Cov_{\omega}(x_k, x_j)$ en (44) que implican a *k*. Bajo esta regla la contribución del factor *k* a la desigualdad global viene dada por

$$\begin{aligned} \overline{Var_{\omega}(x_k)} &= Var_{\omega}(x_k) + \sum_{j \neq k} Cov_{\omega}(x_k, x_j) \\ &= \sum_{j} Cov_{\omega}(x_k, x_j) = Cov_{\omega}(x_k, x) \end{aligned} \qquad (50)$$

y obviamente la suma de todas estas contribuciones para cada uno de los factores nos proporciona $Var_{\omega}(x)$. Esto es lo que Shorrocks (1982) llama la descomposición "natural" de la varianza. En este caso la contribución porcentual de cada factor a la desigualdad global viene dada por el cociente entre la contribución del factor *k*, (50), y la medida global de desigualdad, $Var_{\omega}(x)$, es decir

$$\frac{\overline{Var_\omega(x_k)}}{Var_\omega(x)} = \frac{Cov_\omega(x_k, x)}{Var_\omega(x)} \tag{51}$$

contribuciones que naturalmente suman la unidad. Es importante observar que tales contribuciones pueden tomar valores nulos o negativos, cuando $Cov_\omega(x_k, x) \leq 0$; y que valores negativos indican que el factor *k* actúa como compensador de las diferencias en renta derivadas de los otros factores. Por analogía con la descomposición "natural" del índice de Gini el término $\overline{Var_\omega(x_k)}$ puede ser denominado **pseudo-varianza** para el factor *k*, computacionalmente dicho término puede ser obtenido como la suma de los elementos de una fila (columna) de la matríz de covarianzas entre los factores que componen la renta *per capita* agregada.

Como mencionamos en el epígrafe anterior la **varianza** no es una medida usual de desigualdad, ya que no satisface la propiedad de independencia respecto a la escala, en su lugar se suelen utilizar la **varianza de los logaritmos** o el **cuadrado del coeficiente de variación**. La varianza de los logaritmos no admite una descomposición "natural" en el sentido utilizado más arriba, ya que el logaritmo de una suma no es igual a la suma de los logaritmos; las propiedades de los logaritmos serán utilizadas, no obstante, en el siguiente epígrafe en el que consideraremos la desagregación de la renta *per capita* en factores multiplicativos. Sin embargo las mismas consideraciones derivadas para la varianza se aplican al **cuadrado del coeficiente de variación**, que si satisface la propiedad de independencia respecto a la escala. En este caso

$$\begin{aligned} CV_\omega(x)^2 &= \frac{Var_\omega\left(\Sigma_{k=1}^{K} x_k\right)}{\mu^2} \\ &= \sum_k \frac{Cov_\omega(x_k, x)}{\mu^2} = \sum_k \frac{\overline{Var_\omega(x_k)}}{\mu^2} = \sum_k \overline{CV_\omega(x_k)^2} \end{aligned} \tag{52}$$

Por lo tanto $\frac{\overline{Var_\omega(x_k)}}{\mu^2} = \overline{CV_\omega(x_k)^2}$, que podemos denominar **pseudo-coeficiente de variación** al cuadrado, constituye la elección obvia de la contribución del factor *k* a la desigualdad global.[38] De esta forma, cuando los diversos factores que componen la renta presentan correlación no nula, los términos de interacción son asignados en la misma forma en que se

[38] Obsérvese, sin embargo, que

$$CV_\omega(x_k)^2 = \frac{Var_\omega(x_k)}{\mu_k^2} \quad \text{es diferente de} \quad \frac{Var_\omega(x_k)}{\mu^2} \quad \text{ya que } \mu_k \neq \mu \ \forall k .$$

hizo para la varianza, con lo que obtenemos de nuevo lo que Shorrocks (1982) llama la descomposición "natural" del **cuadrado del coeficiente de variación**. La propiedad más relevante de esta descomposición es que la contribución porcentual de cada factor a la desigualdad global es la misma que la que obteníamos para la descomposición "natural" de la varianza, en efecto, $\frac{\overline{CV_{\omega}(x_k)^2}}{CV_{\omega}(x)^2} = \frac{Cov_{\omega}(x_k, x)}{Var_{\omega}(x)} = \frac{\overline{Var_{\omega}(x_k)}}{Var_{\omega}(x)}$, por lo que en términos de contribuciones porcentuales es irrelevante que hablemos de la varianza o del cuadrado del coeficiente de variación.

Estas descomposiciones "naturales", que se derivan de la propia forma funcional del índice en cuestión,[39] se sustentan sobre una regla arbitraria de asignación de los efectos interacción, otras reglas podrían determinar otras descomposiciones igualmente válidas y quizás más atractivas según en que contextos. Ello nos lleva a la conclusión de que, **en general**, **las contribuciones porcentuales de los distintos factores a la desigualdad no están determinadas de forma única**. De hecho si imponemos solamente dos restriciones débiles, (i) que los factores deben ser tratados de forma simétrica, es decir, que intercambiar las distribuciones de dos factores cualesquiera implica intercambiar también sus contribuciones a la desigualdad global, y (ii) que la contribución de cada factor a la desigualdad global debe ser independiente del nivel de desagregación de los restantes factores; es posible demostrar (Shorrocks (1982)) que una gran parte de los índices utilizados habitualmente en la literatura pueden ser descompuestos de tal forma que la contribución a la desigualdad global atribuida a cada factor es idéntica que la que se obtendría de la descomposición "natural" de cualquier otro índice; en otras palabras la contribución a la desigualdad global de los diferentes factores que componen la renta agregada no está univocamente determinada. Este resultado es en cierta forma descorazonador y requiere cierta cautela en la interpretación de las descomposiciones anteriores, que pueden sér útiles en ciertos contextos pero totalmente estériles en otros.

Una restricción más fuerte que las dos anteriores, que impone que la contribución de dos factores a la desigualdad global sea la misma solamente si (i) la distribución de la renta procedente de ambos factores en idéntica y (ii) si ambos factores suman el total de la renta, permite a Shorrocks (1982) demostrar que bajo estas condiciones hay una regla de descomposición única para una gran parte de los índices de desigualdad utilizados habitualmente, y en la que la contribución porcentual de cada factor a la desigualdad global viene dada por la proporción obtenida en la descomposición "natural" de la varianza, o lo que es lo mismo en la descomposición "natural" del cuadrado del coeficiente de variación. Por

[39] Shorrocks (1982) obtiene también descomposiciones "naturales" para la **clase de medidas generalizadas de entropía** o **familia de índices de Theil**.

tanto, si estamos dispuestos a asumir esta restricción adicional, la contribución porcentual de cada factor a la desigualdad global será independiente del índice de desigualdad elegido y será además facilmente computable.

Abordamos finalmente, al igual que en el epígrafe anterior, cuestiones de **interpretación**, ya que la frase "contribución del factor *k* a la desigualdad global" puede interpretarse, como argumenta Shorrocks (1982), de dos formas diferentes:

(i) Por una parte podemos estar interesados en responder a la pregunta: **¿Cual sería el nivel de desigualdad global observado si el único origen en las diferencias de renta se debiera al factor *k*?**. La respuesta a esta pregunta sugiere aplicar un índice de desigualdad, digamos *I*, a la distribución hipotética de la renta *per capita* en la que la distribución del factor *k* permanece inalterada, pero para el resto de factores se supone una distribución igualitaria; es decir $I_{1k} = I(\mu - \mu_k + x_k)$. En este caso la "contribución del factor *k* a la desigualdad global" será elevada si el valor de I_{1k} es elevado en términos relativos respecto al valor de $I(x)$, obsérvese que en general no hay ninguna razón por la cual I_{1k} deba ser menor que *I*.

(ii) Alternativamente podemos estar interesados en responder a la pregunta: **¿en cuanto se reduciría la desigualdad global si las diferencias de renta en el factor *k* fueran eliminadas, manteniéndose la misma distribución para el resto de factores?**. La respuesta a esta pregunta sugiere comparar el índice de desigualdad global, $I(x)$, con el que se obtendría al aplicar dicho índice a la distribución hipotética de la renta *per capita* en la que la distribución del factor *k* se supusiera igualitaria, mientras que para el resto de factores su distribución se mantuviera inalterada, $I(x - x_k + \mu_k)$; de esta forma la reducción en la desigualdad global vendría dada por $I_{2k} = I(x) - I(x - x_k + \mu_k)$. En este caso la "contribución del factor *k* a la desigualdad global" será elevada si la igualación de la renta procedente de dicho factor reduce de forma significativa el índice global de desigualdad. Obsérvese que I_{2k} puede ser negativo, en cuyo caso, el factor *k* tiende a compensar por las diferencias de renta derivadas de los otros factores.

Shorrocks (1982) muestra que en general no puede establecerse una relación clara entre la "contribución del factor *k* a la desigualdad global" obtenida a partir de las descomposiciones anteriores y los términos I_{1k} e I_{2k} que dan respuesta a las preguntas intuitivas que nos acabamos de formular. Este resultado negativo sugiere que **la descomposición factorial estudiada en este epígrafe sólo será satisfactoria para un conjunto limitado de índices de desigualdad**. La naturaleza del problema puede ser mejor entendida si consideramos la **varianza** como medida de desigualdad, en este caso

$$I_{1k} = Var_{\omega}(\mu - \mu_k + x_k) = Var_{\omega}(x_k) \tag{53}$$

$$\begin{aligned} I_{2k} &= Var_{\omega}(x) - Var_{\omega}(x - x_k + \mu_k) \\ &= Var_{\omega}(x_k) + 2\sum_{j \neq k} Cov_{\omega}(x_j, x_k) \end{aligned} \tag{54}$$

por tanto la interpretación (i), (53), genera la contribución directa o “pura” del factor k a la desigualdad global, e ignora todos los efectos interacción; por su parte la interpretación (ii), (54), añade a dicha contribución directa todos los efectos interacción que implican al factor k y los asigna a la contribución de dicho factor.[40] Obsérvese que, en general, $I_{1k} \neq I_{2k}$ y $\Sigma_k I_{1k} \neq Var_{\omega}(x) \neq \Sigma_k I_{2k}$, por lo que ninguna de las dos interpretaciones intuitivas que hemos considerado anteriormente genera una descomposición consistente,[41] excepto cuando los diferentes factores de renta están incorrelacionados. Sin embargo para el caso de la **varianza** se verifica la siguiente relación entre la “contribución del factor k a la desigualdad global” dada por la descomposición “natural”, (50), y las interpretaciones (i) y (ii)

$$\overline{Var_{\omega}(x_k)} = Cov_{\omega}(x_k, x) = \frac{1}{2}(I_{1k} + I_{2k}) \tag{55}$$

De forma similar, y dado que los experimentos conceptuales propugnados por estas interpretaciones no alteran μ, la relación (55) se verifica tambien para el **cuadrado del coeficiente de variación**, ya que

$$I_{1k} = CV_{\omega}(\mu - \mu_k + x_k)^2 = \frac{Var_{\omega}(x_k)}{\mu^2} \tag{56}$$

$$\begin{aligned} I_{2k} &= CV_{\omega}(x)^2 - CV_{\omega}(x - x_k + \mu_k)^2 \\ &= \frac{Var_{\omega}(x_k) + 2\Sigma_{j \neq k} Cov_{\omega}(x_j, x_k)}{\mu^2} \end{aligned} \tag{57}$$

por tanto

[40] En términos de la matriz de covarianzas de los diferentes factores de renta, x_k, (54) es equivalente a sumar la fila y la columna k y restarle la varianza, para no incurrir en una doble contabilización de la misma.

[41] Entendemos por **descomposición consistente** aquella en la que las contribuciones de los diferentes factores suman el índice de desigualdad global.

$$\overline{CV_{\omega}(x_k)^2} = \frac{Cov_{\omega}(x_k, x)}{\mu^2} = \frac{1}{2}(I_{1k} + I_{2k}) \tag{58}$$

Así pues para estas dos medidas, para las que es posible obtener una descomposición consistente, existe una relación sencilla entre la “contribución de un factor a la desiguald global” y las dos interpretaciones intuitivas consideradas. Sin embargo dado que esta relación no existe en general puede resultar interesante en ocasiones obtener los índices $I(\mu - \mu_k + x_k)$ e $I(x - x_k + \mu_k)$ y compararlos directamente con $I(x)$ para un índice de desigualdad cualquiera, lo que puede proporcionarnos intutición adicional.

Desagregación de la renta per capita en factores multiplicativos.

El interés por el tipo de descomponibilidad estudiado en esta sección proviene de la observación de que la renta *per capita* puede escribirse en ocasiones como el producto de una serie de factores, por ejemplo si nuestras unidades económicas de partida no son individuos, sino agrupaciones de individuos, tales como países o regiones, entonces la renta *per capita* de un espacio geográfico determinado puede descomponerse en el producto de los siguientes factores[42]

$$x_i = \frac{Y_i}{N_i} = \frac{Y_i}{E_i} \cdot \frac{E_i}{PA_i} \cdot \frac{PA_i}{PET_i} \cdot \frac{PET_i}{N_i} \tag{59}$$

donde Y_i renta total de la agrupación i,

E_i empleo total de la agrupación i,

PA_i población activa total de la agrupación i,

PET_i población en edad de trabajar de la agrupación i, y

N_i población total de la agrupación i;

de nuevo resulta de interés preguntarse si es posible analizar la contribución a la desigualdad global de cada uno de los factores que aparecen en la desagregación multiplicativa (59). Este tipo de descomposiciones no han sido objeto de atención por parte de la literatura teórica sobre índices de desigualdad, a pesar de lo cual existen numerosas aplicaciones de algunos de los resultados que mencionaremos en el ámbito del análisis regional (Cuadrado (1991), Esteban (1994), Villaverde (1996, 1997), Goerlich y Mas (1998a)). Ello significa que las propiedades de las descomposiciones que vamos a analizar son desconocidas, especialmente en el ámbito de

[42] Consideraremos cuatro factores que son los naturales en el caso que nos ocupa, pero extensiones a un número mayor o menor son inmediatas.

la interpretación; de hecho una de las contribuciones de este trabajo es llamar la atención sobre problemas interpretativos de algunas descomposiciones sencillas que se obtienen gracias al álgebra de los logaritmos.[43]

Llamando, $pr_i = \frac{Y_i}{E_i}$ a la productividad, $e_i = \frac{E_i}{PA_i}$ a la tasa de empleo, $ta_i = \frac{PA_i}{PET_i}$ a la tasa de actividad, y $d_i = \frac{PET_i}{N_i}$ a la estructura demográfica, podemos escribir $x_i = pr_i.e_i.ta_i.d_i$. Resulta interesante observar como dicha desagregación multiplicativa en la que estamos interesados, (59), adopta una estructura peculiar, que impone ciertas restricciones en el análisis, y que puede no verificarse en otros casos de interés.

En general consideraremos que la renta *per capita* x_i puede escribirse como el producto de k factores

$$x_i = \prod_{k=1}^{K} w_{ki} \tag{60}$$

donde w_{ki} es el factor multiplicativo k en la obtención de la renta *per capita*, x_i.

En el caso de nuestro ejemplo observamos además que las medias para cada una de nuestras variables viene dada por

$$pr = \frac{Y}{E} = \sum_i \frac{E_i}{E}.pr_i \qquad e = \frac{E}{PA} = \sum_i \frac{PA_i}{PA}.e_i$$

$$ta = \frac{PA}{PET} = \sum_i \frac{PET_i}{PET}.ta_i \qquad d = \frac{PET}{N} = \sum_i \frac{N_i}{N}.d_i$$

por lo que podemos escribir $\mu = \frac{Y}{N} = \frac{Y}{E}.\frac{E}{PA}.\frac{PA}{PET}.\frac{PET}{N} = pr.e.ta.d$.

En general consideraremos que la renta *per capita* media, μ, puede ser escrita como el producto de las medias de los k factores considerados

$$\mu = \prod_{k=1}^{K} \eta_k \tag{61}$$

[43] Un análisis más detallado del ofrecido aquí se encuentra en Goerlich (1998a).

donde η_k es la media aritmética del factor w_k, $\eta_k = \Sigma_{i=1}^{n} q_{ki} w_{ki}$ siendo q_{ki} la frecuencia relativa asociada a w_{ki}.

En los casos particulares que vamos a analizar ahora es importante que distingamos entre la "verdadera" media aritmética de w_k, η_k, y la media de w_k evaluada bajo la distribución de x, es decir $\overline{\eta}_k = \Sigma_{i=1}^{n} p_i w_{ki}$, sólo cuando las frecuencias relativas asociadas al factor w_k y a la renta *per capita*, x, coincidan, $q_{ki} = p_i \ \forall i$, entonces $\overline{\eta}_k = \eta_k$, pero en general no hay ninguna razón para suponer tal coincidencia por lo que normalmente $\overline{\eta}_k \neq \eta_k$. Esta observación es especialmente importante ya que, al contrario de lo que sucedía en la desagregación aditiva (43) en la que las ponderaciones de cada factor eran idénticas a las del agregado, esto no sucede en el caso de la desagregación multiplicativa (59), en la que las ponderaciones de cada factor son diferentes y en general diferentes tambien de las frecuencias relativas asociadas al agregado.

Dadas las propiedades de los logaritmos la factorización (60) puede descomponerse aditivamente como, $\log x_i = \Sigma_{k=1}^{K} \log w_{ki}$, e igualmente para μ a partir de (61), $\log \mu = \Sigma_{k=1}^{K} \log \eta_k$. Ello implica que, aparentemente, los resultados del epígrafe anterior, relativos a la desagregación de la renta en factores aditivos, son aplicables, con pequeñas modificaciones, a aquellos índices de desigualdad que utilicen la transformación logarítmica. Sin embargo, el tipo de descomposiciones analizadas en el epígrafe anterior no son operativas en este caso, ya que la frecuencia relativa asociada a cada factor en (60) será, normalmente, diferente y distinta, en general, de la del agregado; por esta razón todo intento de descomponer aditivamente $Var_{\omega}(\log x)$ o $\nu_{\omega}(\log x)$ tropieza con el problema de que la aplicación mecánica de los razonamientos de la sección anterior acaba evaluando los momentos de los factores en (60) bajo la distribución de x, lo que constituye en definitiva un error de especificación. No hay forma aparente de escapar de este problema si insistimos en obtener descomposiciones de los índices de carácter aditivo (Goerlich (1998a)).

El mismo tipo de problema lo encontramos si tratamos de descomponer $T(1)$ o $T(0)$ a partir de la observación

$$\log\left(\frac{x_i}{\mu}\right) = \sum_k \log\left(\frac{w_{ki}}{\eta_k}\right) \tag{62}$$

Para $T(1)$ obtenemos

$$T(1) = \sum_i y_i \log\left(\frac{x_i}{\mu}\right) = \sum_k\left[\sum_i y_i \log\left(\frac{w_{ki}}{\eta_k}\right)\right] = \sum_k \overline{T(1, w_k)} \tag{63}$$

donde $\overline{T(1, w_k)} = \sum_i y_i \log\left(\frac{w_{ki}}{\eta_k}\right)$, es un **pseudo-índice de Theil *T*(1)** para el factor k y que tomaremos como la contribución del factor k a la desigualdad global.

Igualmente para $T(0)$

$$T(0) = -\sum_i p_i \log\left(\frac{x_i}{\mu}\right) = \sum_k\left[-\sum_i p_i \log\left(\frac{w_{ki}}{\eta_k}\right)\right] = \sum_k \overline{T(0, w_k)} \tag{64}$$

donde $\overline{T(0, w_k)} = -\sum_i p_i \log\left(\frac{w_{ki}}{\eta_k}\right)$ es un **pseudo-índice de Theil *T*(0)** para el factor k y que tomaremos como la contribución del factor k a la desigualdad global.

Obsérvese, sin embargo, que las ponderaciones utilizadas en la construcción de $\overline{T(1, w_k)}$, y_i, y en la de $\overline{T(0, w_k)}$, p_i, no aparecen, normalmente, en la propia definición de w_k, lo que hace que estos índices no sean verdaderos índices de Theil; de hecho estos pseudo-índices, al contrario de lo que sucede con $T(1)$ o $T(0)$, no son no-negativos por construcción.

Ya mencionamos anteriormente como una medida de desigualdad es básicamente una media aritmética ponderada de funciones de distancia entre puntos, donde la ponderación es la frecuencia relativa. El problema es que no existe una forma única de ponderar, en el contexto de la renta *per capita* resulta lógico ponderar por la estructura demográfica o por las proporciones de renta, sin embargo no hay ningún argumento *a priori* que nos impida utilizar otras ponderaciones como frecuencias relativas. En el contexto de la desagregación (63) la medida de distancia entre w_{ki} y su media, η_k, es ponderada por las proporciones de renta, y_i, aunque esta variable no aparezca en la propia definición de w_k. Por contra, en el contexto de la desagregación (64) la medida de distancia entre w_{ki} y su media, η_k, es ponderada por las proporciones de población, p_i, aunque esta variable no aparezca en la propia definición de w_k. Hasta que punto es adecuado utilizar unas u otras proporciones como frecuencias relativas es una cuestión abierta, pero en el contexto de la desagregación (59) y de la propia definición de $T(1)$ y $T(0)$ si estas proporciones son o no adecuadas puede evaluarse por medio de la correlación entre y_i, p_i y las proporciones derivadas de la definición de w_{ki}, sólo una correlación elevada justificaría, siquiera de forma burda, la utilización de alguna de estas proporciones. En cualquier caso esta regla aproximada puede ser de utilidad para conocer que descomposición, (63) ó (64), puede ser la más adecuada en un caso concreto.

Obsérvese que, salvo casos particulares, $\overline{T(1,w_k)}$ no coincidirá con el índice $T(1)$ para la variable w_k, ya que $\overline{T(1,w_k)}$ pondera la desigualdad en la variable w_k por proporciones de renta, y_i, mientras que la aplicación mecánica del índice $T(1)$ a la variable w_k la ponderaría por las proporciones derivadas del numerador en la definición de dicha variable; estas proporciones coinciden para la productividad, primer factor considerado en (59), pero no para el resto de variables analizadas. Igualmente, salvo casos particulares, $\overline{T(0,w_k)}$ no coincide con el índice $T(0)$ para la variable w_k, ya que $\overline{T(0,w_k)}$ pondera la desigualdad en la variable w_k por proporciones de población, p_i, mientras que la aplicación mecánica del índice $T(0)$ a la variable w_k la ponderaría por las proporciones derivadas del denominador en la definición de dicha variable; estas proporciones coinciden para la estructura demográfica, último factor considerado en (59), pero no para el resto de variables analizadas. Así por ejemplo $T(0)$ para la productividad, $pr_i = Y_i / E_i$, ponderaría la desigualdad en dicha variable, $\log(pr_i / pr)$, por la estructura del empleo, E_i/E, mientras que $\overline{T(0,pr)}$ la pondera por la estructura demográfica, p_i.

Las desagregaciones (63), $T(1) = \sum_k \overline{T(1,w_k)}$, y (64), $T(0) = \sum_k \overline{T(0,w_k)}$, nos permiten determinar que parte de la desigualdad observada global, $T(1)/T(0)$, se deben a desigualdades en productividad, $\overline{T(1,pr)} / \overline{T(0,pr)}$, desigualdades en las tasas de empleo, $\overline{T(1,e)} / \overline{T(0,e)}$, desigualdades en las tasas de actividad, $\overline{T(1,ta)} / \overline{T(0,ta)}$, y desigualdades en la estructura demográfica, $\overline{T(1,d)} / \overline{T(0,d)}$, tomando como frecuencias relativas siempre las proporciones de renta en el caso de $T(1,w_k)$ o la estructura demográfica en el caso de $T(0,w_k)$. Resulta interesante observar como la importancia relativa de cada factor en dichas descomposiciones es independiente del nivel de desagregación de los restantes factores considerados, lo que constituye una propiedad claramente deseable de toda descomposición aditiva.

Abordamos finalmente, al igual que en los epígrafes anteriores, cuestiones de **interpretación**; de nuevo la frase "contribución del factor k a la desigualdad global" puede interpretarse de dos formas diferentes:

(i) Por una parte podemos estar interesados en responder a la pregunta: **¿Cual sería el nivel de desigualdad global observado si el único origen en las diferencias de renta se**

debiera al factor *k*?. La respuesta a esta pregunta sugiere aplicar un índice de desigualdad, digamos *I*, a la distribución hipotética de la renta *per capita* en la que la distribución del factor *k* permanece inalterada, pero para el resto de factores se supone una distribución igualitaria; es decir $I_{1k} = I\left(\frac{\mu}{\eta_k} w_k\right) = I\left(w_k . \prod_{j \neq k} \eta_j\right)$. En este caso la "contribución del factor *k* a la desigualdad global" será elevada si el valor de I_{1k} es elevado en términos relativos respecto al valor de $I(x)$; en general no hay ninguna razón por la cual I_{1k} deba ser menor que *Y*.

Obsérvese que este experimento conceptual altera la frecuencia relativa asociada a la nueva distribución de la renta *per capita*,[44] ya que $\mu = \prod_j \eta_j = \sum_i q_{ki} w_{ki} . \prod_{j \neq k} \eta_j$, por tanto las nuevas frecuencias relativas son q_{ki}, es decir aquellas que se derivan del factor que origina las desigualdades.

En este caso, por tanto, $I_{1k} = I\left(w_k . \prod_{j \neq k} \eta_j\right)$ no es más que un índice de desigualdad aplicado a la variable $w_k . \prod_{j \neq k} \eta_j$ y utilizando correctamente las ponderaciones derivadas de la definición de la variable w_k. Así pues para cualquier índice que verifique **el principio de independencia respecto a la escala**, la "contribución del factor *k* a la desigualdad global" puede ser medida simplemente mediante la comparación de los índices de desigualdad para *x* y para w_k.

(ii) Alternativamente, y desde un punto de vista general, podemos estar interesados en responder a la pregunta: **¿en cuanto se reduciría la desigualdad global si las diferencias de renta en el factor *k* fueran eliminadas, manteniéndose la misma distribución para el resto de factores?**. La respuesta a esta pregunta sugiere comparar el índice de desigualdad global, $I(x)$, con el que se obtendría al aplicar dicho índice a la distribución hipotética de la renta *per capita* en la que la distribución del factor *k* se supusiera igualitaria, mientras que para el resto de factores su distribución se mantuviera inalterada, $I\left(x \frac{\eta_k}{w_k}\right) = I\left(\eta_k \prod_{j \neq k} w_j\right)$; de esta forma la reducción en la desigualdad global vendría dada por $I_{2k} = I(x) - I\left(\eta_k \prod_{j \neq k} w_j\right)$. En este caso la "contribución del factor *k* a la desigualdad global" será elevada si la igualación de la renta procedente de dicho factor reduce de forma significativa el índice global de desigualdad.

[44] Al menos si queremos mantener intacta, como parece razonable, la media de la distribución de la renta *per capita*.

Obsérvese que I_{2k} puede ser negativo, en cuyo caso, el factor k tiende a compensar por las diferencias de renta derivadas de los otros factores.

De nuevo este experimento conceptual altera la frecuencia relativa asociada a la nueva distribución de la renta *per capita*, sin embargo, al contrario de lo que sucedía con la interpretación anterior dichas frecuencias no se pueden obtener ahora con total generalidad y tampoco tienen un sentido claro.[45] Esto hace que esta interpretación no sea, salvo casos particulares, operativa.

Los comentarios anteriores muestran que en general no puede establecerse una relación clara entre la "contribución del factor k a la desigualdad global" obtenida a partir de las descomposiciones anteriores y los términos I_{1k} e I_{2k} que tratan de dar respuesta a las preguntas intuitivas que nos acabamos de formular. Este resultado negativo, todavía más contundente que en el caso de la descomposición a partir de una desagregación de la renta en factores aditivos, sugiere que **la descomposición factorial estudiada en esta sección es en gran medida estéril.** De hecho parte de los experimentos conceptuales que nos hemos planteado son lógicamente inconsistentes, así por ejemplo no parece posible determinar cual sería el nivel de desigualdad global observado en renta *per capita* si no existieran diferencias en las tasas de empleo entre agrupaciones, ya que igualar las tasas de empleo implica transferencias de ocupados y/o población activa, lo que claramente afecta a otros factores en la desagregación (59).

Estas puntualizaciones sugieren que **el tipo de cuestiones que hemos tratado de analizar en este epígrafe sólo son útiles para un conjunto limitado de preguntas**, pero no en general. En consecuencia las descomposiciones que hemos ofrecido esta sección deben ser interpretadas con cuidado, en concreto los resultados cuantitativos derivados de (63) y (64) deben tomarse a nivel meramente indicativo y con las debidas precauciones. En cualquier caso una actitud razonable sugiere la obtención de los índices de desigualdad para cada uno de los factores, w_k, y su comparación directa con el de la renta *per capita* agregada, $I(x)$, para un índice de desigualdad dado. La descomposición que acabamos de analizar exige pues una actitud de mirar a todas partes.

Finalmente señalar que, al igual que sucedía con la descomposición por subgrupos de población, las descomposiciones consideradas en esta sección y la anterior son puramente descriptivas y no están basadas en consideraciones estadísticas.

[45] El lector puede comprobar esta afirmación experimentando con la descomposición factorial multiplicativa dada por (59).

4. A MODO DE CONCLUSIONES

Estamos ahora en condiciones de ofrecer unas breves conclusiones.

- Desde el punto de vista teórico la cuestión de interés **es como podemos selecionar una medida de desigualdad en una situación particular**. La respuesta depende, al menos parcialmente, del contexto en el que nos movamos y debe ser planteada como un proceso en dos etapas, debemos distinguir entre (i) elegir entre una familia de medidas de desigualdad, y (ii) elegir un miembro particular de esa familia.

- A nivel puramente descriptivo, y en una primera aproximación, los estadísticos tradicionales son útiles, pero su alcance es limitado. Si nos centramos en el análisis agregado, y más concretamente en la problemática de la convergencia entre unidades económicas, parece conveniente utilizar otros estadísticos de dispersión que puedan ayudarnos a caracterizar la distribución de la variable de interés, ya sea renta *per capita* o productividad. En cualquier caso debe observarse que los estadísticos utilizados son ponderados por la frecuencia relativa, estructura demográfica en nuestro ejemplo, tal y como postula la lógica estadística, pero sin embargo ello no es la práctica habitual en la literatura macroeconómica sobre convergencia (Goerlich y Mas (1998c)). Si estamos más interesados por problemas intrínsecos a lo que se denomina desigualdad en el ámbito microeconómico entonces debemos mantener una actitud de mirar a todas partes, en esta situación resulta conveniente que el índice elegido dependa de un parámetro al que poder atribuir significado en términos de nuestra aversión a la desigualdad, o dicho de otra forma de un parámetro que nos permita atribuir un peso diferente a diferentes partes en la distribución de la variable de interés, la **familia de índices de Theil o Atkinson** posee claras ventajas es esta situación. No obstante el examen de proporciones, mencionado en conexión con la curva de Lorenz, se vuelve relevante dado el carácter multidimensional de la desigualdad. Esta forma no concluyente de examinar la desigualdad puede resultar muy útil a la hora de interpretar la evolución de un índice concreto.

- Con respecto a la elección de un miembro particular de una familia de índices el factor clave a tener en cuenta es la sensibilidad del índice a la información en diferentes partes de la distribución, en este sentido la elección de β o ε depende de dos consideraciones: (1) nuestra aversión particular a la desigualdad, y (2) la capacidad discriminatoria del índice en cuestión. El principal punto a tener en cuenta en relación a este segundo argumento es que si se elige un valor del parámetro que represente una gran aversión a la desigualdad, la mayoría de distribuciones que encontremos indicarán un elevado nivel de desigualdad, de forma que será dificil determinar que distribución particular es más desigual que otra. En la práctica valores de β alrededor del entorno [0, 1] o de ε en el entorno de 1 parecen ser razonables.

- Si deseamos analizar el origen de las desigualdades entonces la propiedad de descomponibilidad se muestra clave en el problema de selección de un índice concreto.
  En lo que hace referencia a la descomponibilidad por subgrupos de población la familia de índices de Theil en general y $T(0)$ en particular parece ser el candidato obvio.
  Si nos centramos en la descomponibilidad a partir de una desagregación aditiva de la renta entonces el cuadrado del coeficiente de variación parece ser el índice más apropiado, si bien los resultados de Shorrocks (1982) muestran que bajo ciertas condiciones restrictivas la contribución porcentual de cada factor a la desigualdad global será independiente del índice de desigualdad elegido. No obstante para el análisis de los efectos de la imposición y la medición de la equidad de un sistema redistributivo las pseudo-curvas de Lorenz y los pseudo-índices de Gini asociados a ellas son los más apropiados
  Finalmente si estamos interesados en la descomponibilidad a partir de una desagregación multiplicativa de la renta *per capita* entonces no parece que exista un índice adecuado. La simplicidad con la que es posible descomponer índices tales como $T(1)$ o $T(0)$ se ve oscurecida por problemas interpretativos por lo que debe prevalecer una actitud de mirar a todas partes, es decir los índices para cada uno de los factores deben ser examinados en relación con los de la renta *per capita*. Además si alguno de los factores considerados es una tasa porcentual será conveniente examinar medidas de desigualdad absolutas, tales como la varianza, y no sólo relativas.

REFERENCIAS BIBLIOGRÁFICAS

Aitchison, J. & Brown, J. A. C. (1957) *The lognormal distribution*, Cambridge University Press, Cambridge.

Atkinson, A. B. (1970) "On the measurement of inequality", *Journal of Economic Theory*, 3, 244-263.

Atkinson, A. B. (1980) "Horizontal equity and the distribution of the tax burden" in H. J. Aaron & M. J. Boskins (Eds.) *The Economics of Taxation*, The Brookings Institution, Washington, D. C., Chp.-1, p.-3-18.

Barro, R. J. (1991) "Economic growth in a cross section of countries", *Quarterly Journal of Economics*, 106, (May), 407-443.

Barro, R. J. & Sala-i-Martin, X. (1991) "Convergence across states and regions", *Brookings Papers on Economic Activity*, 1, (April), 107-182.

Barro, R. J. & Sala-i-Martin, X. (1992) "Convergence", *Journal of Political Economy*, 100, 2, 223-251.

Barro, R. J. & Sala-i-Martin, X. (1995) *Economic Growth*, McGraw Hill, New York.

Baumol, W. J. (1986) "Productivity growth, convergence, and welfare", *American Economic Review*, 76, 5, (December), 1072-1085.

Beach, C. M. & Davidson, R. (1983) "Distribution-free statistical inference with Lorenz curves and income shares", *Review of Economic Studies*, 50, 723-764.

Beach, C. M. & Kaliski, S. F. (1986) "Lorenz curve inference with sample weights: An application to the distribution of unemployment experience", *Applied Statistics*, 35, 1, 38-45.

Bentham, J. (1907) *An Introduction to the Principle of Morals and Legislation*, Oxford University Press, Oxford.

Bishop, J. A.; Chakraborti, S. & Thistle, P. D. (1994) "Relative inequality, absolute inequality, and welfare: Large sample tests for partial orders", *Bulletin of Economic Research*, 46, 1, 41-59.

Bishop, J.; Chow, K. V. & Formby, J. P. (1994) "Testing for marginal changes in income distributions with Lorenz and concentration curves", *International Economic Review,* 35, 479-488.

Bishop, J.; Formby, J. P. & Thistle, P. (1992) "Convergence of the south and the non-south income distributions, 1969-79", *American Economic Review*, 82, 262-272.

Blackorby, C. & Donalson, D. (1978) "Measures of relative inequality and their meaning in terms of social welfare", *Journal of Economic Theory*, 18, 59-80.

Blackorby, C. & Donalson, D. (1980) "Ethical indices for the measurement of poverty", *Econometrica*, 48, 1053-1060.

Bosch, A.; Escribano, C. & Sánchez, I. (1989) *Evolución de la Desigualdad y la Pobreza en España. Estudio Basado en las Encuestas de Presupuestos Familiares 1973-74 y 1980-81.* Instituto Nacional de Estadística. (Febrero).

Chakravarty, S. R. (1983) "Ethical flexible measures of poverty", *Canadian Journal of Economics*, 16, 74-85.

Chakravarty, S. R. (1990) *Ethical Social Index Numbers*, Springer Verlag, Berlin.

Chakravarty, S. R. (1997) "On Shorrocks' reinvestigation of the Sen poverty index", *Econometrica*, 65, 5, (September), 1241-1242.

Cowell, F. (1995) *Measuring Inequality*, 2nd Edition, LSE Handbooks in Economics, Prentice Hall, London. (1st. Edition 1977, Phillip Alan Publiserhs Limited, London).

Cowell, F. & Mehta, F. (1982) "The estimation and interpolation of inequality measures", *Review of Economic Studies*, 49, 273-290.

Cuadrado Roura, J. R. (1991) "Las disparidades regionales en la Comunidad Europea y en España", *De Economía Pública*, 12, 3, 107-122.

Dalton, H. (1920) "The measurement of inequality of income", *Economic Journal*, 30, 348-361.

David, H. A. (1968) "Gini's mean difference rediscovered", *Biometrika*, 55, 573.

Davidson, R. & Duclos, J.-Y. (1997) "Statistical inference for the measurement of the incidence of taxes and transfers", *Econometrica*, 65, 6, (November), 1453-1465.

Davies, J. B. & Shorrocks, A. F. (1978) "Assesing the quantitative importance of inheritance in the distribution of wealth", *Oxford Economic Papers*, 30, 138-149.

Deaton, A. & Muellbauer, J. (1980) *Economics and Consumer Behavior*, Cambridge University Press, Cambridge.

DeLong, J. B. (1988) "Productivity growth, convergence, and welfare: A comment", *American Economic Review*, 78, 5, (December), 1138-1155.

del Río, C. & Ruiz-Castillo, J. (1996) "Ordenaciones de bienestar e inferencia estadística. El caso de las EPF de 1980-81 y 1990-91", Segundo Simposio sobre la desigualdad de la renta y la riqueza, *La Desigualdad de Recursos*, Vol.-VI, Fundación Argentaria, Madrid, 9-44.

Díaz-Gimenez, J.; Quadrini, V. & Ríos-Rull, J.-V. (1997) "Dimensions of inequality: Facts on the U.S. distributions of earnings, income, and wealth", *Federal Reserve Bank of Minneapolis, Quarterly Review*, (Spring), 3-21.

Duclos, J.-Y. (1993) "Progressivity, redistribution and equity, with application to the british tax and benefit system", *Public Finance/Finances Publiques*, 48, 350-365.

Esteban, J. M. (1994) "La desigualdad interregional en Europa y en España: Descripción y análisis", en Esteban, J.M. & Vives, X. (Eds.) *Crecimiento y Convergencia Regional en España y en Europa*, 2 volúmenes, Vol 2, Cap.-1, 13-84.

Esteban, J. M. (1996) “Desigualdad y polarización. Una aplicación a la distribución interprovincial de la renta en España”, *Revista de Economía Aplicada*, 4, 11, (Otoño), 5-26.

Esteban, J. M. & Ray, D. (1993) “El concepto de polarización y su medición”, en *Igualdad y Distribución de la Renta y la Riqueza*, vol.-2, Fundación Argentaria, Madrid, 1-35.

Esteban, J. M. & Ray, D. (1994) “On the measurement of polarization”, *Econometrica*, 62, 819-852.

Fei, J. C. H.; Ranis, G. & Kuo, S. W. Y. (1978) “Growth and the family distribution of income by factor components”, *Quarterly Journal of Economics*, 92, 17-53.

Gail, M. H. & Gastwirth, J. L. (1978) “A scale-free goodness-of-fit test for the exponential distribution based on the Lorenz curves”, *Journal of the American Statistical Association,* 73, 787-793.

Gastwirth, J. L. & Gail, M. H. (1985) “Simple asymptotically distribution-free methods for comparing Lorenz curves and Gini indices obtained from complete data”, in Basmann, R. L. & Rhodes, G. F. Jr. (Eds.) *Advances in Econometrics*, Vol.-4, Greenwich, JAI Press.

Gini, C. (1912) “Variabilità e mutabilità, contributo allo studio delle distribuzioni e relazioni statistiche”, *Studi Economico-Giuridici dell’ Universiti di Cagliari*, 3, part 2, 1-158.

Goerlich, F. J. (1998a) “Dinámica de la distribución provincial de la renta. I: Un enfoque desde la óptica de la desigualdad”, *Quadern de Treball 69 (nova època),* Facultat de Ciències Econòmiques y Empresarials, Universidad de Valencia.

Goerlich, F. J. (1998b) “Dinámica de la distribución provincial de la renta, 1955-1995: Un enfoque desde la óptica de la desigualdad”, pendiente de publicación en *Revista de Estudios Regionales*.

Goerlich, F. J. & Mas, M. (1997) “Desigualdad, nivel de renta y crecimiento económico: Una nota con datos de la EPF”, Mimeo, IVIE y Universidad de Valencia.

Goerlich, F. J. & Mas, M. (1998a) “Desigualdad y convergencia en el área de la OCDE”, Documento de Trabajo WP-EC 98/09, (Abril), I.V.I.E.

Goerlich, F. J. & Mas, M. (1998b) “Japan/USA: The (apparent) miracle of convergence”, Mimeo, IVIE y Universidad de Valencia.

Goerlich, F. J. & Mas, M. (1998c) “Medición de las desigualdades: Variables, indicadores y resultados”, pendiente de publicación en *Moneda y Crédito*.

Helmert, F. R. (1876) “Die Genauigkeit der Formel von Peters zur Berechnung des wahrscheinlichen Beobachtungsfehlers direkter Beobachtungen gleicher Genauigkeit”, *Astronomische Nachrichten*, 88, nº 2096.

Howes, S. (1993) “Asymptotic properties of four fundamental curves of distributional analysis”, Unpublised, STICERD, London School of Economics.

Kakwani, N. C. (1977) “Measurement of tax progresivity: An international comparison”, *Economic Journal*, 87, 345, (March), 71-80.

Kakwani, N. C. & Podder, N. (1973) "On the estimation of the Lorenz curve from grouped observations", *International Economic Review*, 14, 278-292.

Kakwani, N. C. & Podder, N. (1976) "Efficient estimation of the Lorenz curve and associated inequality measures from grouped observations", *Econometrica*, (January).

Kendall, M. G. & Stuart, A. (1963) *The Advanced Theory of Statistics. Volume 1: Distribution Theory*. 2ª Ed. Griffin, London.

Kolm, S. C. (1969) "The optimal production of social justice", in J. Margolis & H. Guitton (Eds.) *Public Economics*, MacMillan, New York.

Kolm, S. C. (1976a) "Unequal inequalities. I", *Journal of Economic Theory*, 12, 416-422.

Kolm, S. C. (1976b) "Unequal inequalities. II", *Journal of Economic Theory*, 13, 82-111.

Lerman, R. Y. & Yitzhaki, S. (1995) "Changing ranks and the inequality impacts of taxes and transfers", *National Tax Journal*, 48, 45-59.

Lorenz, M. C. (1905) "Methods of measuring the concentration of wealth", *Publications of the American Statistical Association*, 9, 209-219.

Love, R. & Wolfson, M. C. (1976) *Income Inequality: Statistical Methodology and Canadian Illustrations*. Ottawa, Statistics Canada.

Mas, M.; Maudos, J.; Pérez, F. & Uriel, E. (1995) "Growth and convergence in the Spanish provinces", in Amstrong & Vickerman (Eds.) *Convergence and Divergence among European Regions*. Chp.-13, Ed. Pion.

Milanovic, B. (1997) "A simple way to calculate the Gini coefficient and some implications", *Economics Letters*, 56, 45-49.

Mills, J. A. & Zandvakili, S. (1997) "Statistical inference via bootstrapping for measures of inequality", *Journal of Applied Econometrics*, (forthcoming).

Moyes, P. (1989) "Equiproportionate growth of incomes and after-tax inequality", *Bulletin of Economic Research*, 41, 287-293.

Pigou, A. C. (1912) *The Economic of Welfare*, London. (Editado por MacMillan, New York en 1952).

Plotnick, R. (1981) "A measure of horizontal inequity", *Review of Economics and Statistics*, 63, 1, (Feb), 283-288.

Plotnick, R. (1982) "The concept and measurement of horizontal inequity", *Journal of Public Economics*, 17, 373-391.

Quadrini, V. & Ríos-Rull, J.-V. (1997) "Understanding the U.S. distribution of wealth", *Federal Reserve Bank of Minneapolis, Quarterly Review*, (Spring), 22-36.

Quah, D. (1990) "International patterns of growth: I. Persistence in cross-country disparities", Mimeo. Economics Department, MIT. Cambridge, MA.

Quah, D. (1993a) "Galton's fallacy and test of the convergence hypothesis", *The Sandinavian Journal of Economics*, 95, 4, (December), 427-443.

Quah, D. (1993b) "Empirical cross-section dynamics in economic growth", *European Economic Review*, 37, 2/3, (April), 426-434.

Quah, D. (1996a) "Twin peaks: Growth and convergence in models of distribution dinamics", *Economic Journal*, 106, 437, (July), 1045-1055.

Quah, D. (1996b) "Ideas determining convergence clubs", Working Paper, Economics Department, LSE. (April).

Quah, D. (1996c) "Convergence empirics across economies with (some) capital mobility", *Journal of Economic Growth*, 1, (March), 95-124.

Quah, D. (1997) "Empirics for growth and distribution: Stratification, polarization, and convergence clubs", *Journal of Economic Growth*, 2, (March), 27-59.

Rabadan, I. & Salas, R. (1996) "Convergencia y redistribución intertemporal en España: Efecto de los impuestos directos, cotizaciones sociales y transferencias", *Economía Pública*, (Septiembre), Fundación BBV.

Rao, C. R. (1973) *Linear Statistical Inference and its Applications*, 2nd. Edition, John Wiley & Sons, New York.

Rao, V. M. (1969) "Two decompositions of concentration ratio", *Journal of Royal Statistical Society,* 132, 418-425.

Rawls, J. (1971) *A Theory of Justice*, Oxford University Press, Oxford.

Ruiz-Castillo, J. (1986) "Problemas conceptuales en la medición de la desigualdad", *Hacienda Pública Española*, 101, 17-31.

Ruiz-Castillo, J. (1987) "La medición de la pobreza y la desigualdad en España, 1980-1981", Estudios Económicos nº 42, Banco de España, Madrid.

Ruiz-Castillo, J. (1993) "La distribución del gasto en España de 1973-74 a 1980-81", en Almunia, J. y Gutiérrez, L. (Eds.), Primer Simposio sobre igualdad y distribución de la renta y la riqueza, *La Distribución de la Renta*, Vol.-II, Fundación Argentaria, Madrid, 51-89.

Ruiz-Castillo, J. (1997) "A simplified model for social welfare analysis. An aplication to Spain, 1973-74 to 1980-81", Working Paper 97-37, Economic Series 15, (May), Departamento de Economía, Universidad Carlos III de Madrid.

Sala-i-Martin, X. (1994) "Cross-sectional regressions and the empirics of economic growth", *European Economic Review*, 38, 739-747.

Saunders, T. J. (1970) (Traductor y editor) *Plato: The laws*, Penguin, Harmondsworth.

Schutz, R. R. (1951) "On the measurement of income inequality", *American Economic Review*, 41, (March), 107-122.

Sen, A. (1973) *On Economic Inequality*, Oxford University Press, Oxford.

Sen, A. (1976) "Poverty: An ordinal approach to measurement", *Econometrica*, 44, 219-231.

Shorrocks, A. F. (1980) "The class of additively descomposable inequality measures", *Econometrica*, 48, 613-625.

Shorrocks, A. F. (1982) "Inequality decomposition by factor components", *Econometrica*, 50, 193-211.

Shorrocks, A. F. (1984) "Inequality decomposition by population subgroups", *Econometrica*, 52, 1369-1386.

Shorrocks, A. F. (1995) "Revisiting the Sen poverty index", *Econometrica*, 63, 1225-1230.

Theil, H. (1967) *Economics and Information Theory*, Amsterdam, North-Holland.

Villaverde Castro, J. (1996) "Desigualdades provinciales en España, 1955-1991", *Revista de Estudios Regionales*, 45, 89-108.

Villaverde Castro, J. (1997) "Convergencia regional y unión monetaria. ¿Dónde estamos?. ¿A dónde vamos?, *Lecciones* 1/97, Apertura del curso académico 1997/98, Universidad de Cantabria.

Zaiger, D. (1983) "Inequalities for the Gini coefficient of composite populations", *Journal of Mathematical Economics*, 12.

Zubiri, I. (1985) "Una introducción al problema de la medición de la desigualdad", *Hacienda Pública Española*, 95, 291-317.