DESIGUALDAD DE LA DISTRIBUCIÓN DE LA RENTA EN LAS PROVINCIAS DE CASTILLA Y LEÓN: UN ESTUDIO REALIZADO CON MÉTODOS NO PARAMÉTRICOS.

Rafael HERRERIAS PLEGUEZUELO
Federico PALACIOS GONZALEZ
José CALLEJON CESPEDES
Eduardo PEREZ RODRIGUEZ
Facultad de Ciencias Económicas y Empresariales
Universidad de Granada

RESUMEN

Utilizando los estimadores de regresión no paramétrica, que se construyen mediante lo que se conoce como una función núcleo (kernel), se ajustan, para cada provincia y para el conjunto de la Comunidad, curvas de cuantiles de población y renta sobre los datos muestrales (valores de renta recogidos en la EPF 1990 - 1991). Mediante dichas curvas se buscan los respectivos cuantiles necesarios para el cálculo, en cada caso, de la curva de Zenga.

PALABRAS CLAVE: *Curva de Zenga, concentración, regresión no paramétrica, método del kernel.*

1.- INTRODUCCIÓN.

El conocimiento de la desigualdad de la distribución de la renta entre las distintas provincias de una Comunidad Autónoma, como la de Castilla y León, siempre es un objetivo deseable a la hora de proponer las acciones de política económica encaminadas a incrementar el nivel de bienestar y la equidistribución de la renta generada en la Comunidad.

En este sentido, Herrerías, Palacios y Callejón (1996), tomando como valores empíricos de renta, para la Comunidad Autónoma de Castilla y León, los recogidos en la EPF (1990 - 1991) del INE, ajustan un modelo probabilístico para la distribución de la renta, en cada una de las provincias así como para el conjunto de la Comunidad; en este trabajo se utiliza una función generadora polinómica, cuyos parámetros se estiman siguiendo la metodología de Pearson. En trabajos posteriores, Prieto (1998) realiza una modelización paramétrica de la distribución de la renta en España mediante métodos robustos; Pérez, de Prada, Prieto, Rueda y Zarzosa (1998) analizan los factores condicionantes de la desigualdad de la renta en Castilla y León, mediante un análisis de la varianza multifactorial.

Una buena herramienta para el estudio de la concentración sobre los distintos niveles de renta es la curva de concentración de Zenga (1984) que, además de tener una definición relativamente sencilla, permite conocer el mencionado grado de concentración sobre los distintos niveles de renta. En el presente trabajo se obtiene una estimación de la curva de Zenga para cada una de las provincias de la Comunidad de Castilla y León y otra para el conjunto de la Comunidad.

La curva de concentración de Zenga

Para obtener esta curva es necesario calcular cuantiles sobre los porcentajes acumulados de población y renta respectivamente.

Sea F(x) la función acumulativa de proporciones (porcentajes) de población sobre los distintos niveles de renta para una población determinada.

Se define la función acumulativa de proporciones (porcentajes) de renta también sobre los diferentes niveles de renta R(x) de la siguiente forma:

$$R(x) = \frac{\int_0^x tdF(t)}{\int_0^{+\infty} tdF(t)} = \frac{1}{E[x]}\int_0^x tdF(t) \qquad (1)$$

Si x_p e y_p representan los p-cuantiles en las respectivas funciones acumuladas de población y renta respectivamente, es decir si $F(x_p)= p$ y $R(y_p)=p$, se define la curva de Zenga como una función de la proporción p de la siguiente forma (Zenga 1984)

$$z(p) = \frac{y_p - x_p}{y_p} = 1 - \frac{x_p}{y_p} \qquad (2)$$

Cuando esta curva se calcula mediante un modelo de función acumulativa predeterminado, estimando previamente sus parámetros, se corre el riesgo de incorporar a la curva de Zenga cualidades ajenas a los datos muestrales y que tan sólo forman parte del modelo que, en su rigidez, es incapaz de adaptarse totalmente a dichos datos.

El proceso de cálculo de los porcentajes de renta a partir de dicho modelo ajustado y la obtención de cocientes de cuantiles de población y renta para llegar a la curva de Zenga produce un efecto amplificador de la falta de flexibilidad del modelo ajustado. El resultado final puede ser una curva de Zenga con algunas propiedades que no se manifiestan en los datos y pueden conducir a valoraciones sesgadas del índice de concentración. La curva de Zenga, obtenida mediante modelos ajustados de función acumulativa puede presentar crecimientos, que se deben exclusivamente al modelo y que no se hacen patentes en su versión empírica (Salvaterra 1990) obtenida punto a punto, directamente desde los datos. (Véase gráfico 1).

En el presente trabajo se propone el cálculo de la curva de Zenga evitando un modelo que imprima propiedades ajenas a la información muestral. En primer lugar se realiza el cálculo de la curva empírica mediante las técnicas elementales de obtención de cuantiles, propias de la estadística descriptiva y muy efectivas cuando el tamaño de la muestra es elevado. Posteriormente se utilizan métodos de regresión no paramétrica para suavizar las curvas de porcentajes acumulados de población y renta, de las que posteriormente se obtendrá la curva de Zenga.

Curva empírica de Zenga

Esta curva se obtiene punto a punto para cualquier valor $0 < p < 1$, a partir de los cuantiles que se calculan mediante interpolación sobre las proporciones acumuladas de población y renta en las observaciones muestrales:

Sea una muestra x_1 ; x_2 ; ....x_n de rentas obtenidas por individuos de un colectivo y sea $x_{(1)}$; $x_{(2)}$; .....$x_{(k)}$ los k valores de renta distintos y ordenados crecientemente presentes en la muestra anterior. Se define la proporción acumulada de población sobre $x_{(i)}$ como

$$F_i = \frac{1}{n} N\left[x_j \leq x_{(i)}\right]$$

donde $N\left[x_j \leq x_{(i)}\right]$ representa el número de valores muestrales inferiores o iguales $x_{(i)}$

Análogamente la proporción de renta en $x_{(i)}$ como

$$R_i = \frac{\sum_{x_j \leq x_{(i)}} x_j}{\sum_{i=1}^{n} x_i}$$

Para un valor $0 < p < 1$ arbitrario se pueden encontrar, mediante interpolación lineal (Montiel, 1997), los p-cuantiles de población x_p y renta y_p , necesarios para sustituir en (2).

Curva modelizada de Zenga

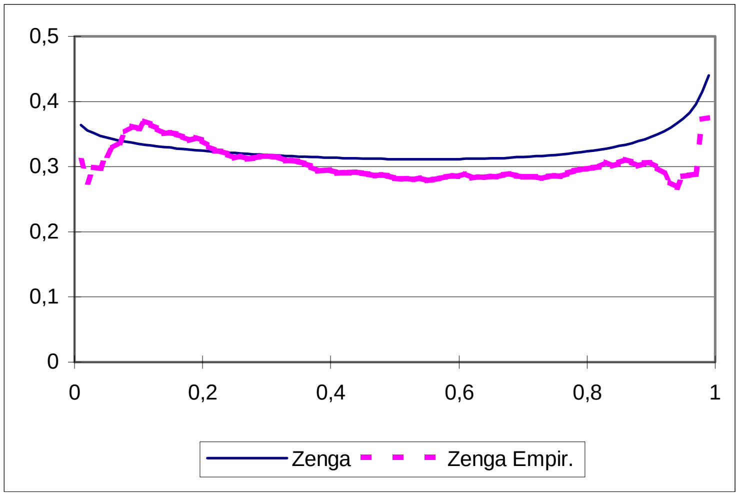
Es habitual ajustar un modelo de probabilidad a los datos de la renta y, a partir de él, obtener la curva de Zenga. Si F(x) representa el modelo ajustado para la población y R(x) el que obtenido para la renta, la curva de Zenga tiene como ecuación

$$z(p) = 1 - \frac{F^{-1}(p)}{R^{-1}(p)}$$

pero esta curva viene caracterizada por las propiedades intrínsecas de F y de R y sobre todo de las resultantes de relacionar ambas mediante un cociente en el que confluyen una estimación de F y una función R obtenida de dicha estimación mediante (1). Sobre este cociente se puede amplificar la falta de flexibilidad del modelo F y sobre todo de R para adaptarse a los datos muestrales. Esto puede provocar crecimientos demasiado acentuados; desplazamiento del máximo del cociente (mínimo de z(p)) o cualquier otro efecto debido exclusivamente a la relación por cociente de los modelos ajustados.

La rigidez del modelo, sobre todo en los extremos, puede apreciarse en el siguiente gráfico. En él, a partir de los datos correspondientes a toda la Comunidad de Castilla y León, se muestra la curva empírica de concentración de Zenga, así como la obtenida para la población a partir de un modelo de Dagum tipo I, ajustado utilizando una metodología similar a la seguida por Fernández, Haro, y Martín, (1996).

Gráfico 1.



Si se quieren evitar los problemas anteriormente descritos, para obtener una curva de Zenga suavizada, parece lógico utilizar una metodología que permita, utilizando modelos de distribución flexibles, ajustar a partir de las proporciones muestrales acumuladas, sendos modelos de funciones acumulativas de población y renta por separado para posteriormente calcular los respectivos p-cuantiles.

Se trata de evitar que especificaciones de modelos, que individualmente resultan medianamente apropiados para las curvas acumulativas, sean inaceptables cuando confluyen sobre la curva de Zenga.

Como caso particular, dentro de esta vía podrían estimarse ambas funciones acumulativas mediante métodos no paramétricos de regresión donde no existe modelo que encorsete los datos y donde el grado de flexibilidad viene regulado por el valor de una constante de suavizamiento que se determina a partir de la propia muestra.

2.- MÉTODO DEL NÚCLEO Y OBTENCIÓN DE LA CURVA DE ZENGA

Las funciones de porcentajes (proporciones) acumulados están acotadas, suelen tener curvaturas poco acusadas y pueden suponerse estrictamente crecientes en el rango muestral de la variable renta, a efectos de aplicación práctica al problema que en el presente trabajo nos ocupa. Son pues de fácil estimación mediante métodos de suavizamiento sin incurrir en sesgos apreciables.

Parece oportuno utilizar métodos de regresión no paramétrica que permitan establecer una relación entre cualquier valor p, comprendido entre cero y uno, como variable independiente y los respectivos cuantiles x_p e y_p de población y renta, al objeto de obtener a partir de ellos la curva de Zenga tal y como se establece en (2).

Siguiendo la metodología propuesta en Herrerías, Palacios, Callejón y Pérez (1998), en este trabajo, para un h determinado, se utiliza el estimador de regresión de Nadaraya-Watson con una función núcleo K para estimar los cuantiles de población y renta, dado un valor de p, de la siguiente forma:

$$x_p = \frac{\sum_{i=1}^{k} x_{(i)} K\left(\frac{p-F_i}{h}\right)}{\sum_{i=1}^{k} K\left(\frac{p-F_i}{h}\right)} \qquad y_p = \frac{\sum_{i=1}^{k} x_{(i)} K\left(\frac{p-R_i}{h}\right)}{\sum_{i=1}^{k} K\left(\frac{p-R_i}{h}\right)}$$

El núcleo utilizado en este caso es el de Epanechnikov definido de la siguiente forma (Härdle 1991)

$$K(u) = \begin{cases} \frac{3}{4}\left(1-u^2\right) & si\ |u| \leq 1 \\ 0 \ \ en\ cualquier\ otro\ caso \end{cases}$$

A partir de los porcentajes suavizados de población y renta se construye la curva de Zenga (suavizada) tal y como se indica en (2).

3.- RESULTADOS OBTENIDOS

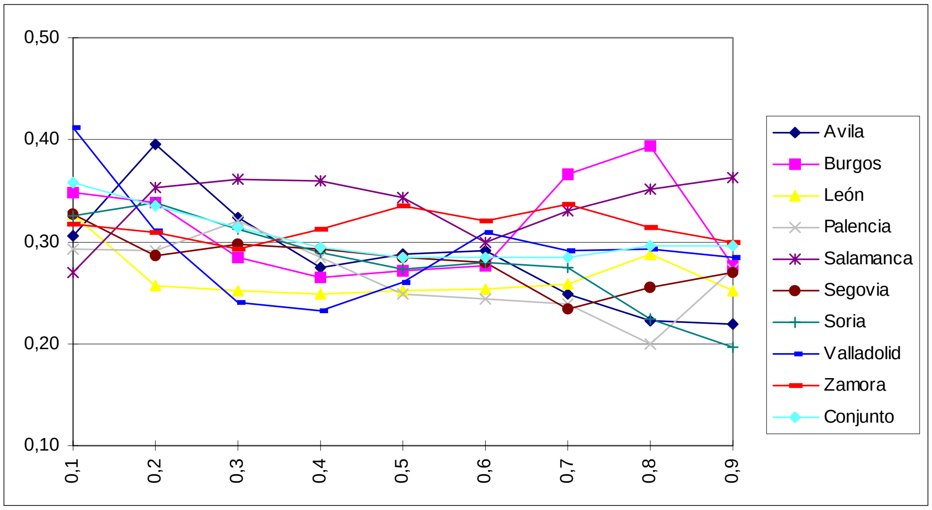
A partir de los valores de renta de la EPF (1990-91), se obtiene, con la metodología expuesta en el epígrafe anterior, la correspondiente curva de Zenga para cada una de las provincias de la Comunidad de Castilla y León y para el conjunto de toda la Comunidad. (Véase ANEXO I).

Las modelizaciones realizadas permiten conocer, en cada caso, la concentración para cada nivel, (considerado en cuantiles), de la renta.

En el siguiente gráfico se ha trazado una línea poligonal para cada provincia utilizando como vértices los valores que toma la curva modelizada de concentración de Zenga, sobre los deciles. Considerando distintos valores de p, se puede observar que no coincide la ordenación, atendiendo al valor que toma la curva de concentración de Zenga, de las provincias de Castilla y León.

Se puede observar, por ejemplo, las sensibles diferencias entre las provincias de Salamanca y Valladolid, cuyas curvas de concentración son tales que una de ellas es casi imagen especular de la otra, mostrando una estructura de distribución de la renta contrapuesta.

También en el gráfico puede observarse que el mayor parecido entre las curvas z(p), de las distintas provincias, se produce en torno al valor de p=0,6.



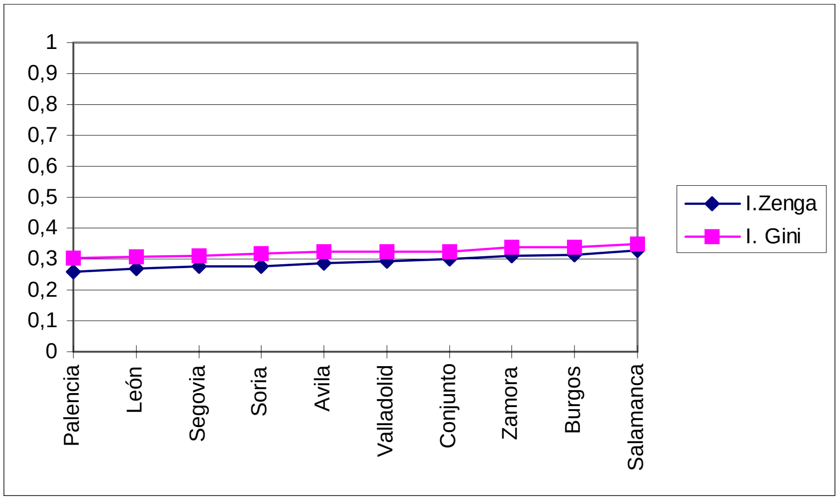
Obsérvese que es posible establecer una ordenación de las distintas provincias para cada uno de los valores de p. En la siguiente tabla se muestra una ordenación de menor a mayor concentración, de las distintas provincias y la Comunidad, para cada uno de los cuantiles considerados: p=0,2; 0,4; 0,6 y 0,8.

Las dos últimas columnas corresponden a la ordenación atendiendo a los índices de Zenga (valor esperado de la curva de concentración, es decir el área bajo la misma) y de Gini respectivamente.

Orden creciente de concentración					
0,2	**0,4**	**0,6**	**0,8**	**Indice Zenga**	**Ind. de Gini**
León	Valladolid	Palencia	Palencia	Palencia	Palencia
Segovia	León	León	Avila	León	León
Palencia	Burgos	Burgos	Soria	Segovia	Segovia
Zamora	Avila	Segovia	Segovia	Soria	Soria
Valladolid	Palencia	Soria	León	Avila	Avila
Conjunto	Soria	Conjunto	Valladolid	Valladolid	Valladolid
Soria	Segovia	Avila	Conjunto	Conjunto	Conjunto
Burgos	Conjunto	Salamanca	Zamora	Zamora	Zamora
Salamanca	Zamora	Valladolid	Salamanca	Burgos	Burgos
Avila	Salamanca	Zamora	Burgos	Salamanca	Salamanca

Si bien la ordenación de las dos últimas columnas de la tabla anterior es la misma, sin embargo los índices de Zenga y de Gini no coinciden pues, en todo caso, este último toma valores ligeramente superiores. Este hecho puede observarse en el siguiente gráfico:

Índices de Gini y de Zenga



CITAS BIBLIOGRÁFICAS

FERNANDEZ, A.; HARO, J. Y MARTÍN, G. Medición de la desigualdad y el bienestar social. Análisis de la distribución de la renta en España. Estudios regionales, Nº 45, (1996), pp 15-42.

HÄRDLE W. Smoothing Tecniques. Springer Verlag. New York. (1991).

HERRERÍAS, R.; PALACIOS, F. Y CALLEJÓN, J. Distribución de la renta en la Comunidad de Castilla y León: dos métodos de estimación. Actas del 5º Congreso de Economía Regional de Castilla y León. Comunicaciones 2, Avila, (1996). pp. 978-990

HERRERÍAS, R.; PALACIOS, F.; CALLEJÓN, J. Y PÉREZ, E. Obtención de la curva de concentración de Zenga por el método del kernel. Actas de la XII Reunión Anual de ASEPELT-ESPAÑA. Córdoba, (1998).

MONTIEL, A.M., RIUS, F. Y BARÓN F.J. Elementos básicos de Estadística Económica y Empresarial. Prentice Hall. Madrid. (1997).

PÉREZ, A; DE PRADA, M.D. PRIETO M. RUEDA, E. Y ZARZOSA P. Factores condicionantes de la desigualdad de la renta en Castilla y León 1990-1991. Trabajo subvencionado por la Consejería de Economía y Hacienda de la Junta de Castilla y León. Valladolid, (1998).

PRIETO, M. Modelización paramétrica de la distribución personal de la renta para España mediante métodos robustos. Tesis Doctoral. Universidad de Valladolid. (1998)

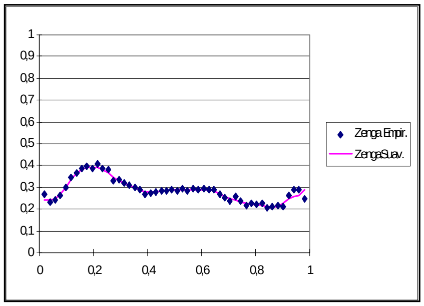
SALVATERRA T.: “Comparisons Among Concentration Curves and Indexes in Some Empirical Distributions", en Dagum, C y Zenga, M, (1990), 194-214.

ZENGA, M. Proposta per un Indice di Concentrazione Basato sui Rapporti fra Quantili di Popolazione e Quantili di Reddito. Giornale degli Economisti e Annali di Economia, XLIII, (1984). Pág. 301-326.

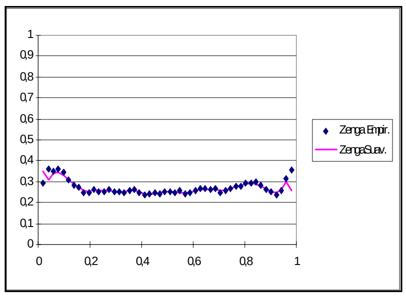
ANEXO I

CURVAS DE CONCENTRACIÓN DE ZENGA

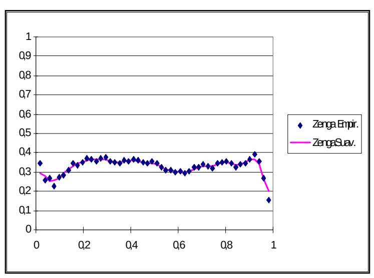
Avila



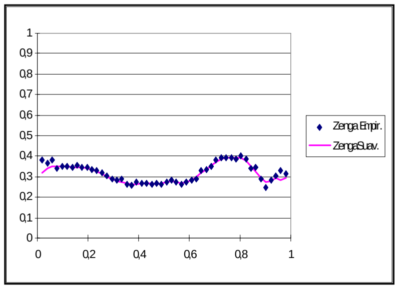
León



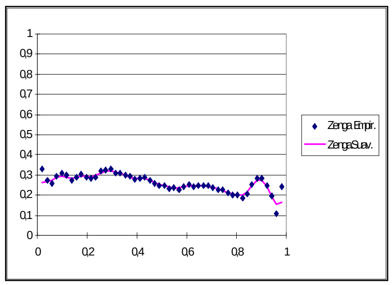
Salamanca



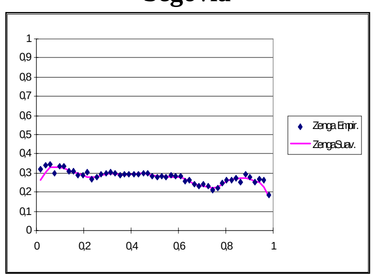
Burgos



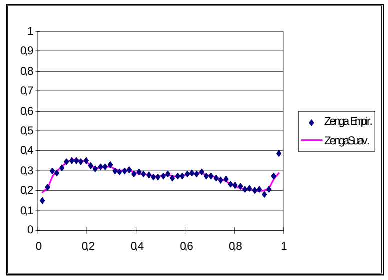
Palencia



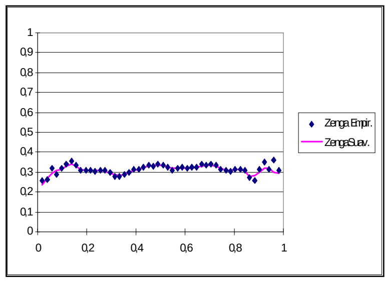
Segovia



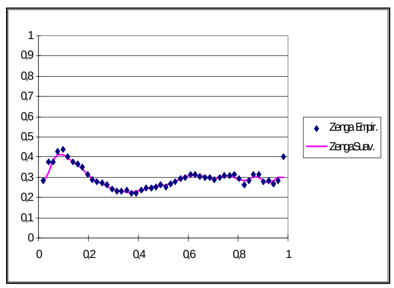
Soria



Zamora



Valladolid



Comunidad de Castilla y León