

**INCIDENCIA DE IMPUESTOS Y PRESTACIONES EN ESPAÑA:
UNA EVALUACIÓN DESDE LA PERSPECTIVA DE LA INFERENCIA
ESTADÍSTICA**

Raquel Andrés*
Samuel Calonge**

*CREB (Centre de Recerca en Economia del Benestar)

**Universitat de Barcelona y CREB

Correspondencia a:

Samuel Calonge

CREB

Baldiri i Reixac 4-6 , 08028 Barcelona

Dpto. Econometría, Estadística y Economía Española
Diagonal 690, 08034 Barcelona

INTRODUCCIÓN

El análisis empírico de la incidencia de impuestos y prestaciones es una herramienta clave para evaluar el efecto redistributivo que ejerce la actuación pública sobre la distribución de la renta. Para el caso español, existe un amplio abanico de estudios en el ámbito de la incidencia y la redistribución, bien desde una óptica parcial, donde se analiza la progresividad de un impuesto(s) o de una prestación(es), o bien desde una óptica conjunta, donde se analiza el efecto neto de impuestos que supone para los hogares la financiación de las transferencias que disfrutan. Respecto al primer tipo de análisis y desde la vertiente de los impuestos, podemos citar entre otros, los trabajos de Lasheras *et al.* (1993), Mayo y Salas (1993), Argimón y González-Páramo (1995), Manresa *et al.* (1996) etc. Por el lado del gasto social, contamos con las referencias de Bandrés (1993), Gimeno (1993), Estruch (1996), Jiménez *et al.* (1993), etc. Un análisis conjunto del efecto de los impuestos y prestaciones que valora el efecto redistributivo neto de la financiación de las prestaciones vía impuestos es el trabajo de Calonge y Manresa (1997, 2001)

Una característica común que comparten estos estudios hace referencia a la naturaleza del análisis, es decir, de base conceptual microeconómica y la información estadística está representada por una muestra de observaciones que corresponden a individuos u hogares. Mediante la imputación de la carga impositiva y de las prestaciones a cada agente económico sujeto a estudio, los resultados que se ofrecen en forma de estimaciones de índices de desigualdad vertical y horizontal, de ordenadas de las curvas de Lorenz o de las curvas de concentración, etc, permiten extraer conclusiones relativas los efectos de progresividad y de redistribución del sistema. Además, otro aspecto relevante en este tipo de trabajos es la comparación dinámica de la desigualdad de la renta o de una distribución impositiva como consecuencias de las políticas económicas que se llevan a cabo en este ámbito (véase Oliver, J *et al.* (2001))

Desde un punto de vista de la metodología estadística aplicada en este tipo de estudios, las estimaciones de los estadísticos se realizan en un contexto de ausencia de inferencia, es decir, no existe información relativa a la variabilidad muestral de los estimadores (error estándar, intervalos de confianza), ni tampoco se efectúan contrastes

de hipótesis sobre los parámetros relevantes¹. Sin embargo, cada vez es mayor el consenso sobre la necesidad de acompañar estos cálculos con información relativa a la precisión muestral de los estadísticos si queremos comparar las distribuciones de la renta: "...Los resultados sugieren que la inferencia estadística es esencial incluso con grandes muestras, Mills y Z (1997)"

Las distribuciones teóricas (asintóticas) de los índices han sido profusamente estudiados en la literatura. Cabe citar los trabajos de Beach y Davison (1983), Cowell (1989), Kakwani (1990), Bishop *et al.* (1998), entre otros. Además, Davidson y Duclos (1997) extienden estas distribuciones en el caso de datos dependientes, cuando utilizamos datos de panel o realizamos comparaciones entre una distribución de la renta neta y bruta basada en una misma muestra de hogares o individuos. Ahora bien, se conoce poco de las propiedades de estas distribuciones ante problemas relativos al tamaño de la muestra, presencia de outliers, etc. De hecho, las distribuciones de la renta, de impuestos y de transferencias son asimétricas, con largas colas y fuerte presencia de outliers.

Un método alternativo a la inferencia clásica consiste en la estimación de la distribución empírica de los índices y ordenadas de Lorenz mediante métodos de remuestreo, por ejemplo, mediante la técnica bootstrap. En el contexto que estamos interesados, la primera referencia que encontramos en la literatura es de Mills y Zandvakili (1997), que comparan los intervalos de confianza bootstrap de los índices de Gini y Theil con los intervalos estimados a partir de la aproximación normal asintótica.

El método bootstrap presenta algunas ventajas. En primer lugar capta la distribución empírica a partir de los datos de la muestra, es decir, no utilizamos una aproximación asintótica. En segundo lugar, las fórmulas teóricas, como veremos a continuación, son complejas. Resulta más sencillo las estimaciones obtenidas mediante el método bootstrap. En tercer lugar, un problema añadido reside cuando los datos son dependientes (datos de panel). Tal y como señala Biewen, M (2002) : "...los métodos bootstrap tienen en cuenta la dependencia temporal de las variables sin necesidad de determinar su estructura de covarianzas..". El presente trabajo tiene como objetivo ilustrar la aplicación de estas técnicas a la estimación de índices de desigualdad, curvas de Lorenz e índices de progresividad y redistribución. En el apartado 2 realizamos una descripción de las distribuciones teóricas de los estadísticos y sus expresiones analíticas.

¹ En lo que nuestro conocimiento alcanza y para el caso español, únicamente en la referencia Martín, P () se realiza, por ejemplo, un test de hipótesis para contrastar el impacto de las prestaciones sociales en la distribución de la renta.

El apartado 3 describe, de forma sencilla, las principales características de los métodos bootstrap. En el apartado 4 incluimos un ejercicio de simulación con objeto comparar las propiedades de las distribuciones asintóticas descritas en el apartado 2 respecto a sus correspondientes versiones bootstrap. Finalmente, en el apartado 5 mostramos una aplicación de estas técnicas al análisis de la redistribución y progresividad de los impuestos y prestaciones en España.

La información utilizada es la muestra de fusión estadística EPF90-IRPF90 (véase Calonge y Manresa (2001), Alegre *et al* (2001)). Esta base de datos se obtiene mediante el enlace estadístico efectuado entre la encuesta de presupuestos familiares 1990 y una muestra representativa de los hogares declarantes del impuesto sobre la renta de las personas físicas del año 1990 IRPF90. Esta muestra combina patrones de gastos de los hogares con ingresos fiscales estimados de la fuente tributaria, además de otra información relevante cómo la categoría socioeconómica, composición de la familia, etc. Se trata entonces de calcular para cada hogar el montante de cada uno de los impuestos que contemplamos y las prestaciones que reciben. Para controlar por la diferente composición demográfica de los hogares, la estimación de la renta del hogar, sus pagos impositivos y las prestaciones es en términos equivalentes, es decir, ponderando las variables por la siguiente escala de equivalencia:

$$h = (\text{adultos} + 0,5\text{hijos})^{0,5}$$

donde adultos e hijos representan el número de adultos y el número de hijos en el hogar.

DISTRIBUCIONES ASINTÓTICAS.

En esta apartado se describen los estadísticos y su correspondiente varianza muestral de las siguientes medidas de desigualdad de la renta y de concentración: Gini, Theil(1), ordenadas de la curva de Lorenz y de concentración. Además, se incluyen también las expresiones analíticas de los estadísticos de progresividad de impuestos y prestaciones (Kakwani y Suits) La demostración de la varianza del índice de Kakwani se encuentra en Bishop *et al.* (1998) La obtención de la varianza del índice de Suits es de mayor complejidad y se incluye en este trabajo como aportación original (véase también

Andrés, R (2002) para una descripción más detallada) Las tablas 1 y 2 incluye las expresiones de los estadísticos utilizados y la expresión del estimador de su varianza asintótica. En el próximo apartado describimos el método bootstrap para estimar la varianza o el error estándar de los estadísticos y sortear la dificultad matemática y de programación que comportan estas expresiones.

INSERTAR TABLAS 1 Y 2

EL MÉTODO BOOTSTRAP

El método bootstrap es esencialmente un método de remuestreo para estimar la distribución empírica de un estadístico, por ejemplo, un índice de desigualdad, las ordenadas de la curva de Lorenz, etc que denotamos por I . Consiste básicamente en la extracción de R muestras independientes, de tamaño n , con reemplazamiento, a partir de una muestra original (cada extracción da lugar a una remuestra bootstrap) Para cada remuestra se calcula el estadístico \hat{I}^* , que es la replicación bootstrap del estadístico \hat{I} . Por ejemplo, el gráfico 1 muestra la distribución empírica obtenida a partir de 1000 repeticiones bootstrap del índice de Gini de la renta equivalente antes de impuestos.

INSERTAR GRÁFICO 1

Después de R extracciones, el vector de dimensión R de replicas bootstrap \hat{I}^* se emplea para calcular una estimación de la variabilidad muestral de \hat{I} , por ejemplo, mediante la fórmula de la desviación estándar. En este trabajo estamos interesados en intervalos de confianza y contrastes de hipótesis del tipo $H_0: I = 0$, $H_0: I_1 = I_2$.

El método más sencillo para estimar intervalos de confianza bootstrap es el método percentil. El proceso es el siguiente: a partir de R remuestras bootstrap, obtenemos una estimación de la función de distribución empírica \hat{F}_R del estadístico:

$$\hat{F}_R(x) = \frac{1}{R} \sum_{i=1}^R I_{-\infty, x}(\hat{I}_i^*)$$

Para un nivel de significación α del intervalo de confianza, el método percentil consiste en calcular los percentiles de la distribución bootstrap $\hat{F}_R^{-1}(\frac{\alpha}{2})$ y $\hat{F}_R^{-1}(1 - \frac{\alpha}{2})$.

El intervalo de confianza será entonces:

$$\left[\hat{I}_{L1}^*, \hat{I}_{L2}^* \right] = \left[\hat{F}_R^{-1}\left(\frac{\mathbf{a}}{2}\right), \hat{F}_R^{-1}\left(1 - \frac{\mathbf{a}}{2}\right) \right]$$

Estas expresiones se refieren a una situación en que $R \rightarrow \infty$, pero en la práctica se generan un número determinado de R muestras bootstrap independientes para calcular los intervalos (es recomendable en la construcción de intervalos de confianza una cifra de $R = 1000$) En nuestra aplicación utilizamos una versión mejorada de este método denominada BCa, acrónimo de bootstrap corregido por el sesgo y acelerado ("bias-corrected and accelerated" (en Efron y Tibshirani (1993) encontramos una explicación rigurosa y detallada del método).

El método bootstrap-t, para construir intervalos de confianza de un estadístico está basado en la distribución *t-student* y el intervalo de confianza definido a partir de ésta:

$$t_{(n-1)} = \frac{\hat{I} - I}{\widehat{es}(\hat{I})} \quad \left[\hat{I} - t_{n-1}^{(1-\mathbf{a})} \cdot \hat{es}(\cdot), \hat{I} - t_{n-1}^{(\mathbf{a})} \cdot \hat{es}(\cdot) \right]$$

donde $t_{n-1}^{(\mathbf{a})}$ es el α -percentil de la distribución *t-student* con $n-1$ grados de libertad. La estimación bootstrap-t consiste en calcular:

$$t^*(r) = \frac{\hat{I}^*(r) - I}{\hat{es}^*(r)}$$

donde $\hat{I}^*(r)$ y $\hat{es}^*(r)$ son las replicaciones bootstrap del estadístico de interés y su error estándar. Mediante un total de R réplicas se trata entonces de reproducir la "tabla bootstrap-t" utilizando el método percentil anterior, de tal forma que la estimación del percentil- α es calculado por:

$$\#\{t^*(r) \leq \hat{t}^{(\mathbf{a})}\} / R = \mathbf{a}$$

y el intervalo de confianza bootstrap-t se define por²:

$$\left[\hat{I} - \hat{t}^{(1-\mathbf{a})} \cdot \hat{es}, \hat{I} - \hat{t}^{(\mathbf{a})} \cdot \hat{es} \right]$$

² En el caso de un estadístico complejo para el cual no conocemos una expresión analítica del error estándar, se hace necesario la estimación bootstrap del error estándar para cada remuestra bootstrap. Ello supone dos niveles anidados de remuestreo (remuestrear cada remuestra), lo cual aumenta notablemente la carga computacional del proceso además de añadir un nuevo error de aproximación.

SIMULACIÓN

En este apartado realizamos un ejercicio de simulación con objeto de comparar las propiedades de los estimadores con distribución muestral asintótica conocida con los estimadores bootstrap. Esta simulación se efectúa para los índices de desigualdad, curvas de Lorenz e índices de progresividad propuestos. Se trata entonces de evaluar la precisión de estos métodos con datos de renta e impuestos procedentes de la muestra de fusión, por tanto son datos reales y no simulados a partir de una distribución paramétrica de la renta.

La comparación de los estadísticos de la distribución normal asintótica y las distribuciones bootstrap esta basada en los intervalos de confianza y los niveles de cobertura real del parámetro. Calculamos las siguientes estimaciones: método estándar1: intervalos de confianza construidos con las fórmulas de la aproximación normal asintótica, método estándar 2: formula asintótica pero sustituyendo la expresión del error estándar por su correspondiente estimación bootstrap, Método BCa y bootstrap-t. En general, la realización de un experimento de este tipo no es tan evidente puesto que no existe una distribución de referencia (especialmente en muestras finitas) necesaria para realizar la comparación³.

El experimento propuesto en este apartado está basado en considerar la muestra de fusión, que tiene por tamaño muestral $n = 21\ 497$ observaciones, el universo de referencia. La simulación realizada consiste en efectuar 2000 replicaciones del siguiente algoritmo, para tres tamaños de muestra diferentes $n=200$, $n=1000$, $n=2000$

1. Extracción de una muestra aleatoria simple sin reemplazamiento de tamaño n del universo de la simulación
2. Calcular el estadístico de interés sobre la muestra extraída (gini, theil, coordenadas de Lorenz e índices de progresividad)
3. Generación de $R = 1000$ remuestras bootstrap (extracciones con reemplazamiento del mismo tamaño que la muestra original n) sobre la muestra extraída en (1). Utilizar las remuestras bootstrap para calcular los intervalos de confianza, método estándar1, estándar2, Bca, bootstrap-t
4. Acumular para cada intervalo el valor del indicador de cobertura ($I_c=I$, si intervalo contiene al parámetro), calcular las medidas de forma y longitud, la cobertura real aproximada y los valores medios de forma, longitud y extremos

³ Biewen (2002) realiza un ejercicio de Montecarlo asumiendo una distribución log-normal de la renta.

INSERTAR TABLA 2

Los valores de la tabla 2 de la simulación muestran que la cobertura real de los intervalos de confianza bootstrap son superiores a aquellos obtenidos mediante las fórmulas asintóticas (método estándar 1). Para el índice de Gini y las decilas de la curva de Lorenz, los métodos bootstrap ofrecen estimaciones de la cobertura real de los intervalos similares a su cobertura nominal. Las discrepancias entre la probabilidad de cobertura real y nominal de los intervalos de confianza oscila en un intervalo de medio a dos puntos porcentuales si tenemos en cuenta el tamaño de $n = 2000$. El método percentil corregido, BCa, tiene un comportamiento muy similar al método bootstrap-t. Esto es relevante puesto que BCa tiene un coste computacional más reducido.

Con relación al índice de desigualdad de Theil, ambos métodos presentan niveles de cobertura real alejados de su cobertura nominal, aunque aquí también la precisión de los intervalos bootstrap es superior (por ejemplo, para $n=2000$ la cobertura del método estándar1 es igual a 0.718 mientras que para bootstrap-t la cobertura real es del orden de 0.867) Estos niveles de cobertura real están relacionados obviamente con la tasa de convergencia asintótica de la distribución empírica del índice de Theil a la aproximación normal. Es de interés también destacar las diferencias que muestran los dos tipos de intervalos construidos a partir de las distribuciones muestrales asintóticas de los estadísticos, métodos estándar(1) y estándar(2). Los intervalos relativos al método estándar(2) se obtienen sustituyendo el error estándar asintótico por su estimación bootstrap. Las diferencias entre la cobertura nominal y real de los intervalos alcanza entre 0,16 y 0,06 puntos porcentuales a favor de este último. Con respecto a los índices de progresividad, se reproducen prácticamente los mismos comentarios, es decir, superioridad de los intervalos bootstrap sobre las fórmulas estándar. Los índices de progresividad para los impuestos directos presentan peores resultados de cobertura que los obtenidos para los impuestos indirectos. La cobertura real más próxima su valor nominal corresponde al bootstrap-t, con un valor de 0,925 para el caso del índice de Kakwani y $n = 2000$. La mejor estimación que corresponde al índice de Suits es de 0,899, es decir, seis puntos porcentuales inferior a su valor nominal. BCa presenta valores con una precisión más pequeña, pero las diferencias son notables con los métodos estándar. Por ejemplo, si utilizamos las expresiones del apartado anterior para construir los intervalos de confianza de los índices de Kakwani y Suits de los impuestos directos, los niveles de cobertura real alcanzados son 0,841 y 0,716, muy alejados de sus valores

nominales. En general, la reducción del error de cobertura utilizando los intervalos bootstrap es notorio. En el caso de la imposición indirecta y para ambos índices, los métodos BCa y bootstrap-t proporcionan valores muy próximos a su probabilidad nominal de cobertura, tanto para los tamaños de $n=1000$ y $n=2000$, muy superiores también a los ofrecidos por los métodos estándar. El resultado bootstrap-t incluso es prácticamente exacto a su probabilidad nominal. En general, los métodos bootstrap originan estimaciones con una mayor precisión muestral, y es notoria su diferencia con los intervalos obtenidos mediante la aproximación asintótica.

REDISTRIBUCIÓN Y PROGRESIVIDAD DE IMPUESTOS Y PRESTACIONES

En este apartado realizamos una aplicación de los métodos descritos en los apartados anteriores utilizando la muestra de fusión estadística IPRF90-EPF90. La imputación de la carga impositiva y de las prestaciones para cada hogar se ha realizado en consonancia con las hipótesis de incidencia usuales. El impuesto sobre la renta de las personas físicas ha sido atribuido al hogar cuyos miembros soportan estatutariamente dicho impuesto. También se asigna al hogar los impuestos locales (IBI y vehículos) que pagan, además de las cuotas de la seguridad social de los trabajadores y aquella parte de las cotizaciones empresariales soportada por los trabajadores de acuerdo a las hipótesis de traslación, calculada para cada uno de los miembros del hogar activos. El impuesto sobre el valor añadido y los impuestos especiales se imputan a partir del gasto realizado por el hogar. La parte trasladada a consumo del impuesto de sociedades, de las cuotas pagadas por los empleadores a la Seguridad Social y de los impuestos sobre las importaciones ha sido imputada en proporción al consumo realizado por cada uno de los hogares respecto al total de gasto estimado. La parte del impuesto de sociedades que pagan los propietarios del capital se distribuye entre estos en proporción a sus ingresos que provienen de las rentas del capital (intereses, dividendos, valor de los alquileres imputados a las viviendas) respecto al total de dichas rentas. Las prestaciones han sido asignadas siguiendo el criterio del flujo del beneficio, para un detalle pormenorizado de la incidencia de las prestaciones y su cuantificación (véase Calonge y Manresa (2001) y Calonge 2002) para una descripción detallada de las hipótesis de traslación) El total de impuestos asignados entre los hogares corresponde al 92% del total de ingresos fiscales del año 1990, cifra suficientemente significativa del montante total de impuestos. El sistema fiscal, de acuerdo con las

conclusiones obtenidas en Calonge, S (2002) es ligeramente progresivo (un porcentaje del 2,75% de la carga fiscal global es trasladada de ricos a pobres) y con un efecto redistributivo pequeño (2,3% de la renta después de impuestos es redistribuida) El papel redistributivo descansa entonces en el efecto de las prestaciones sobre la distribución de la renta.

Desde un punto de vista de la inferencia nos preguntamos por la significación estadística de los índices, en otros términos, queremos generalizar los resultados estimados a partir de una muestra a toda la población. Por otro lado, también es importante calcular y contrastar las diferencias de progresividad y redistribución entre impuestos (o prestaciones), es decir, establecer un rango de "dominancia" entre las diferentes figuras impositivas y de las transferencias que reciben las familias.

En lugar de presentar los resultados sobre índices agregados de desigualdad, progresividad, etc, creemos más ilustrativo presentar el análisis redistributivo y de progresividad en términos de resultados de dominancia de las distribuciones de la renta antes de impuestos, de la renta después de impuestos y de las curvas de concentración de impuestos y prestaciones (véase, por ejemplo, Davidson y Duclos (1997)) De esta forma estudiamos los efectos a lo largo de toda la distribución y no mediante un índice sintético. Esto permite analizar y contrastar, sobre quién recae el esfuerzo redistributivo de un determinado impuesto o prestación, cuanta renta se redistribuye entre los distintos agentes, etc. Para ello, definimos el perfil redistributivo de un impuesto o una prestación mediante la distancia entre la curva de concentración de la renta después de impuestos $Y^* = Y - T^d$ y la curva de Lorenz antes de impuestos Y :

$$[IR(p) = L_{Y^*}(p) - L_Y(p)]$$

Esta distancia proporciona información relativa a la distribución de la renta después de impuestos respecto de la distribución obtenida por un impuesto proporcional. La medida $IR(p)$ se interpreta como la fracción de renta después de impuestos que se traslada desde los hogares situados por encima del percentil $100p$ hacia los hogares con rentas inferiores a dicho percentil, como consecuencia de la "no-proporcionalidad". El gráfico 2 muestra el perfil redistributivo del impuesto sobre la renta, de los impuestos directos e indirectos y del total de las prestaciones, así como los intervalos de confianza de las ordenadas con objeto de contrastar el impacto redistributivo y la dominancia entre

⁴ En el caso de una prestación P (efectivo o en especie) , $Y^* = Y + P$.

las curvas $IR(p)$. El contraste de la hipótesis nula $H_0: IR(p)=0$ es un contraste de dominancia de la curva de Lorenz antes de impuestos sobre la curva de concentración después de impuestos, de otra manera, de la existencia de un efecto redistributivo significativo de un impuesto (prestación) como consecuencia de la no-proporcionalidad. Para contrastar la dominancia de una curva A frente a otra B, rechazamos la hipótesis nula de no-dominancia a favor de la alternativa de dominancia si cada punto de la curva A es significativamente mayor que el correspondiente punto de la curva B a un nivel de significación del 5%, es decir, si no observamos “no-solapamiento entre las curvas”

INSERTAR GRÁFICO 2

El área situada por debajo de cada una de las curvas representadas y el eje de abscisa p mide la redistribución total ejercida por el impuesto o la prestación (de hecho este cálculo es equivalente a dos veces el índice de Reynolds-Smolensky, si no existe reordenamiento) Se interpreta como la fracción total de renta después de impuestos o de renta final (en el caso de prestaciones) que se traslada de ricos a pobres debido a la progresividad del sistema. Por otra parte, la parte positiva (o negativa) de las curvas $IR(p)$ revela que proporción de la renta después de impuestos excede (o es inferior) de la que se obtendría mediante una distribución proporcional. Del gráfico se desprenden las siguientes características:

1. El impuesto sobre la renta, los impuestos directos (que incluyen IRPF, cotizaciones del trabajador y empleador soportada por el trabajador, sociedades, otros impuestos locales) y las prestaciones tienen un impacto positivo y estadísticamente significativo sobre la redistribución de la renta. La imposición indirecta opera en sentido contrario, y su efecto es claramente significativo. Los intervalos de confianza no incluyen en ningún caso el eje de abscisa, que define la hipótesis nula de redistribución cero.
2. La curva de concentración del IRPF presenta un perfil pronunciado con un máximo en la octava decila ($IR(p)=0,031$) El impuesto redistribuye un porcentaje del 3,14% de la renta Y^* desde la decilas novena y décima hacia el resto de la población de hogares. El impuesto de sociedades presenta el mismo perfil redistributivo que el IRPF (bajo una hipótesis de incidencia progresiva), aunque

su impacto redistributivo es más pequeño (0,7% de la renta después de impuesto es redistribuida) La curva que corresponde al total de impuesto domina a la curva de IRPF hasta la séptima decila, pero son indistinguibles a partir de aquí, probablemente porque las cotizaciones sociales tienen un papel regresivo en las decilas más altas.

3. Los resultados resaltan la importancia de las prestaciones en la redistribución de la renta. La curva de concentración de las prestaciones exhibe una dominancia clara respecto al resto de las curvas y su efecto redistributivo tiene su máximo en la decila sexta (redistribución del 7% la renta final desde las decilas superiores a la sexta hacia el resto de la población) El gráfico 3 muestra ahora la curva de Lorenz de la renta antes de impuestos y la curva de Lorenz de la renta final (incluyendo prestaciones) con objeto de comparar ambas distribuciones teniendo en cuenta el reordenamiento al pasar de una distribución a otra. El efecto redistributivo de las prestaciones son claramente significativos a lo largo de toda la distribución, lo que refuerza nuestro resultado, en términos de significación estadística, del impacto de las prestaciones sobre la distribución de la renta. Esto corrobora el argumento relativo al papel que tienen las prestaciones como “motor redistributivo” del sistema.

INSERTAR GRÁFICO 3

El gráfico 4 desglosa el efecto redistributivo de las prestaciones a lo largo de la distribución de la renta. Los resultados muestran claramente la dominancia entre las diferentes curvas (puesto que no existe solapamiento entre los intervalos de confianza) En el sistema de prestaciones son las transferencias monetarias a las familias (pensiones contributivas y no contributivas, desempleo) las de mayor impacto redistributivo, aunque las prestaciones en especie (sanidad y educación) son también cualitativamente importante, De hecho, tal y como muestra la parte de la curvas con pendiente positiva, el 60% de la población con menos recursos son beneficiarios, si lo comparamos con la situación de distribución equiproporcional de las prestaciones. Por último, es relevante destacar el efecto redistributivo prácticamente nulo del gasto público universitario, puesto que su curva de concentración está muy próxima al eje de abcisas (no-redistributivo), aunque su efecto es estadísticamente distinto de cero.

INSERTAR GRÁFICO 4

La diferencia $IR(p_i) - IR(p_{i-1})$ identifica el efecto redistributivo de la renta después de impuestos para una decila específica. Habíamos comentado antes que el impuesto sobre la renta de las personas físicas tiene un perfil netamente redistributivo, con un máximo de redistribución en las decila novena, en otros términos, el 3,1% de la renta después de impuestos se redistribuye desde el 20 más rico hacia el resto de la población. Si observamos el gráfico 5, la contribución de la decila décima es, evidentemente, la más importante. En términos porcentuales indica que la redistribución originada por la decila décima es del orden del 87,1% de toda la renta distribuida por este impuesto y el 12,9% restante corresponde a la novena. Cuando consideramos el efecto redistributivo originado por toda la imposición directa, el perfil redistributivo es menos progresivo, puesto que en este caso son decilas contribuyentes desde la sexta hasta la décima. En concreto, la contribución de la décima decila y novena a la redistribución es del orden del 55% y el 33% respectivamente. Las decilas receptoras del efecto redistributivo captan aproximadamente alrededor de un 20-22% del efecto redistributivo cada una, excepto la cuarta decila que recibe un 15% de la renta redistribuida.

INSERTAR GRÁFICO 5

La imposición indirecta origina redistribución regresiva, es decir, las decilas primera a séptima resultan penalizadas, pues obtendrían una renta después de impuestos mayor si la imposición indirecta corresponde a un impuesto proporcional. Las dos decilas mayores resultan “ganadoras” como consecuencia de la regresividad de los impuestos indirectos. En concreto la renta redistribuida hacia las decilas más ricas, como consecuencia de la imposición indirecta (no-proporcional) es del orden del 11,5%, del cual, el 87% es atribuida a la décima decila. De la misma forma que hemos definido el perfil redistributivo mediante la distancia $IR(p)$, definimos ahora el perfil de progresividad de un impuesto de la siguiente manera:

$$[TR(p) = L_{Y;Y}(p) - C_{Y;T}(p)]$$

que representa la diferencia entre la curva de Lorenz antes de impuestos y la curva de concentración impositiva. Esta distancia, que describimos para los impuestos en el

gráfico 6, se interpreta como la fracción del impuesto total que se traslada desde los hogares con renta menor a la percentil 100p hacia la parte superior de la distribución, como consecuencia de la progresividad del sistema.

INSERTAR GRÁFICO 6

El área entre la curva $TR(p)$ y el eje horizontal es el índice de Kakwani. Los resultados que se muestran permiten investigar los perfiles de progresividad en términos de significación estadística y de dominancia entre las curvas. El impuesto sobre la renta tiene un carácter netamente progresivo. Es relevante también comparar los perfiles de progresividad del impuesto de sociedades bajo dos supuestos de incidencia distintos. La variante sociedades² hace referencia es progresiva y la hipótesis de incidencia adoptada es que $2/3$ del impuesto se traslada a los propietarios del capital y $1/3$ al consumo. La hipótesis sociedades¹ supone traslación al consumo, capital y propietarios de bienes inmuebles en la misma proporción. Hay que subrayar pues que la consideración de la progresividad de este impuesto es bastante sensible a la hipótesis de incidencia que adoptemos.

Hemos querido mostrar la importancia de hacer inferencia sobre las distribuciones de renta, impuestos y prestaciones. La mayoría de las técnicas no paramétricas basadas en "bootstrapping" son sencillas de implementar, con un bajo coste computacional y de fácil programación. Ello permite evaluar los efectos o cambios sobre la distribución de la renta que originan los impuestos y prestaciones y sus potenciales reformas, controlando por el error muestral asociado a cualquier estimación.

TABLA 1. MEDIDAS DE DESIGUALDAD/PROGRESIVIDAD Y VARIANZAS ASINTÓTICAS

ÍNDICE	Estadístico	ESTIMADOR DE LA VARIANZA ASINTÓTICA
<i>Ordenada de Lorenz</i>	$\hat{L}(p_i) = \frac{p_i}{\bar{x}} \hat{I}_i, \quad 0 \leq p_i \leq 1$ <p>donde \hat{I}_i es el estimador muestral de la esperanza condicionada</p> $I_i = \int_0^{k(p_i)} xf(x/X \leq k(p_i))dx$ <p>X: renta antes de impuestos</p>	$\hat{\mathbf{s}}_{Lpi} = \frac{p_i}{\bar{x}^2} [\hat{\mathbf{g}}_i^2 + (1-p_i)(\mathbf{k}(p_i) - \hat{I}_i)^2] + \left(\frac{p_i \hat{I}_i}{\bar{x}^2}\right)^2 \mathbf{s}^2 - 2 \left(\frac{p_i \hat{I}_i}{\bar{x}^3}\right) [\hat{\mathbf{g}}_i^2 + (\mathbf{m} - \hat{I}_i)(\mathbf{k}(p_i) - \hat{I}_i)]$ <p>$\mathbf{k}(p_i)$ estimador muestral del p-cuantil de la distribución de la renta,</p> $\hat{I}_i = \frac{\sum_{j=1}^n x_j I_j}{\sum_{j=1}^n I_j} \quad \hat{\mathbf{g}}_i^2 = \frac{\sum_{j=1}^n I_j x_j^2}{\sum_{j=1}^n I_j} - \hat{I}_i^2$ <p>siendo I_i la función indicador tal que $I_i = 1$, si $x_i \leq \mathbf{k}(p_i)$ y $I_i = 0$ en caso contrario</p>
<i>Gini</i>	$\hat{G} = (2/\bar{x}) \hat{\text{cov}}(x, F(x)) = \frac{d}{2\bar{x}},$ <p>donde $F(x)$ es función de distribución</p>	$\mathbf{s}_G^2 = \frac{1}{4\bar{x}^2} [\mathbf{s}_d^2 - 4\hat{G}n \text{cov}(d, \bar{x}) + 4\hat{G}^2 \mathbf{s}^2],$ <p>donde</p> $\mathbf{s}_d^2 = 4 \left[\frac{1}{n} \sum d_i^2 - d^2 \right], \quad \text{cov}(\bar{x}, d) = \frac{2}{n} \left[\frac{1}{n} \sum x_i d_i - \bar{x}d \right] \quad \text{siendo } d = \frac{1}{n} \sum_{i=1}^n d_i$ <p>y $d_i = \frac{n}{n-1} [2(x_i - \bar{x}_i)p_i - (x_i - \bar{x})]$</p>
<i>Theil(1)</i>	$\hat{T}_1 = (1/n) \sum_{i=1}^n (x_i / \bar{x}) \ln(x_i / \bar{x})$	$\mathbf{s}_{T_1}^2 = \frac{1}{\bar{x}^2} [\hat{E}(x \log x)^2 - \hat{v}^2] + \frac{(\bar{x} + \hat{v})^2 \mathbf{s}^2}{\bar{x}^4} - \frac{2(\bar{x} + \hat{v})}{\bar{x}^3} [\hat{E}(x^2 \log x) - \hat{v}\bar{x}]$ <p>donde</p> $\hat{v} = \frac{1}{n} \sum_{i=1}^n x_i \log x_i, \quad \log(\hat{g}) = \frac{1}{n} \sum_{i=1}^n \log x_i; \quad \hat{E}(x \log x)^2 = \frac{1}{n} \sum_{i=1}^n (x_i \log x_i)^2, \quad \hat{E}(x^2 \log x) = \frac{1}{n} \sum_{i=1}^n x_i^2 \log x_i$

<p><i>Kakwani</i></p>	$\hat{\Pi} = \hat{C}_Y - \hat{G}_X = \frac{d_g}{2\bar{y}} - \frac{d}{2\bar{x}},$ <p>donde \hat{G}_X es el gini de la renta X $\hat{C}_Y = \frac{2}{\bar{y}} \hat{\text{cov}}(g(x), F(x))$ es el índice de concentración de $Y=g(X)$</p>	$\mathbf{s}_{\Pi}^2 = \mathbf{s}_G^2 + \mathbf{s}_C^2 - \frac{2}{\bar{x}\bar{y}} \cdot \left[\hat{\text{cov}}(\bar{x}, \bar{y}) \hat{G}_X \hat{C}_Y n + \frac{\hat{\text{cov}}(d, d_g)}{4} n - \frac{\hat{\text{cov}}(\bar{x}, d_g)}{2} \hat{G}_X n - \frac{\hat{\text{cov}}(\bar{y}, d)}{2} n \hat{C}_Y \right]$ <p>donde $\hat{\mathbf{s}}_G$ y $\hat{\mathbf{s}}_C$ son los estimadores de las varianzas de $\sqrt{n}(\hat{G}_X - G_X)$ y $\sqrt{n}(\hat{C}_Y - C_Y)$ respectivamente y</p> $d_g = \frac{1}{n} \sum_{i=1}^n d_{gi} \text{ siendo } d_{gi} = \frac{n}{n-1} [2(g(x_i) - \bar{g}_i) p_i - (g(x_i) - \bar{y})]$
<p><i>Suits</i></p>	$\hat{S} = \hat{CR}_Y - \hat{G}_X = \frac{d_1}{\bar{x}\bar{y}} - \frac{d_2}{\bar{x}} - \frac{d_3}{2\bar{x}}$ <p>donde \hat{G}_X es el gini de la renta X y CR es índice de concentración relativa $\hat{CR}_Y = \frac{2}{\bar{y}} \hat{\text{cov}}(g(x), F_1(x))$,</p> <p>siendo $F_1(x) = \frac{1}{m_X} \int_0^x y dF(y)$ la proporción acumulada de renta total.</p>	$\mathbf{s}_S^2 = \mathbf{s}_{CR_Y}^2 + \mathbf{s}_G^2 - \frac{2}{\bar{x}^2} \left[\mathbf{s}_X^2 \hat{CR}_Y \hat{G}_X + \frac{\mathbf{s}_{XY}}{\bar{y}^2} \hat{\mathbf{q}}_1 \hat{G}_X + \hat{G}_X n \hat{\text{cov}}(\bar{x}, d_2) + n \frac{\hat{\text{cov}}(d_1, d_3)}{2\bar{y}} \right. \\ \left. - \hat{G}_X n \frac{\hat{\text{cov}}(\bar{x}, d_1)}{\bar{y}} - \hat{CR}_Y n \frac{\hat{\text{cov}}(\bar{x}, d_3)}{2} - \hat{\mathbf{q}}_1 n \frac{\hat{\text{cov}}(\bar{y}, d_3)}{2\bar{y}^2} - n \frac{\hat{\text{cov}}(d_2, d_3)}{2} \right]$ <p>donde $\hat{\mathbf{s}}_G$ y $\hat{\mathbf{s}}_{CR}$ es el estimador de la varianza asintótica de $\sqrt{n}(\hat{G}_X - G_X)$ y $\sqrt{n}(\hat{CR}_Y - CR_Y)$ respectivamente,</p> $d_1 = \frac{1}{n} \sum_{i=1}^n d_{1i} \text{ con } d_{1i} = \frac{n}{n-1} x_i [\bar{y} - \bar{g}_i p_i] + g(x_i) \bar{x}_i p_i$ $d_2 = \frac{1}{n} \sum_{i=1}^n d_{2i} \text{ con } d_{2i} = \frac{n}{n-1} [x_i (1 - p_i) + \bar{x}_i p_i] \text{ y } d_3 = \frac{1}{n} \sum_{i=1}^n d_{3i} \text{ con } d_{3i} = \frac{n}{n-1} [2(x_i - \bar{x}_i) p_i - (x_i - \bar{x})]$

TABLA 2. COBERTURA REAL APROXIMADA PARA LOS ÍNDICES DE DESIGUALDAD, CURVA DE LORENZ Y MEDIDAS DE PROGRESIVIDAD.

Indice	<i>n</i> = 200	<i>n</i> =1000	<i>n</i> = 2000	<i>n</i> = 200	<i>n</i> =1000	<i>n</i> = 2000
	Método	Estándar⁽¹⁾			BCa	
<i>Gini</i>	0,783	0,795	0,788	0,9015	0,929	0,931
<i>Theil</i>	0,948	0,779	0,718	0,83	0,852	0,858
<i>Lorenz Deciles</i>						
<i>1</i>	0,805	0,829	0,855	0,844	0,869	0,949
<i>3</i>	0,807	0,814	0,829	0,930	0,938	0,946
<i>5</i>	0,810	0,812		0,93?	0,938?	
<i>7</i>	0,789	0,807	0,790	0,917	0,933	0,947
<i>9</i>	0,768	0,794	0,782	0,881	0,920	0,920
<i>Kakwani (D)</i>	0,839	0,826	0,841	0,892	0,881	0,901
<i>Kakwani (I)</i>	0,803	0,837	0,823	0,881	0,926	0,930
<i>Suits(D)</i>	0,680	0,685	0,716	0,878	0,855	0,855
<i>Suits(I)</i>	0,826	0,846	0,851	0,864	0,924	0,925
	Metodo	Estándar⁽²⁾			Bootstrap-t	
<i>Gini</i>	0,829	0,879	0,884	0,925	0,952	0,948
<i>Theil</i>	0,699	0,755	0,771	0,818	0,866	0,867
<i>Lorenz Deciles</i>						
<i>1</i>	0,865	0,886	0,911	0,823	0,894	0,961
<i>3</i>	0,868	0,899	0,910	0,929	0,950	0,963
<i>5</i>	0,871	0,893		0,934	0,949	
<i>7</i>	0,851	0,881	0,896	0,933	0,948	0,958
<i>9</i>	0,783	0,849	0,870	0,903	0,940	0,938
<i>Kakwani (D)</i>	0,864	0,826	0,855	0,901	0,882	0,925
<i>Kakwani (I)</i>	0,857	0,897	0,913	0,918	0,935	0,948
<i>Suits(D)</i>	0,828	0,809	0,832	0,878	0,883	0,899
<i>Suits(I)</i>	0,848	0,891	0,906	0,906	0,929	0,943

Cobertura nominal: 0.95

GRÁFICO 1

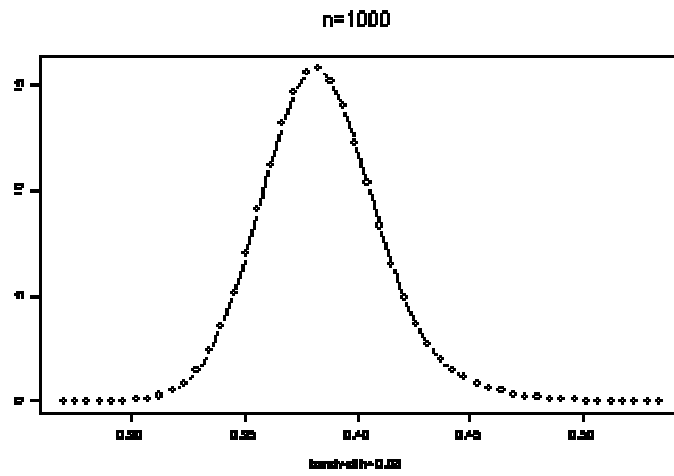


GRÁFICO 2

$$L_y^e(p) - L_y(p)$$

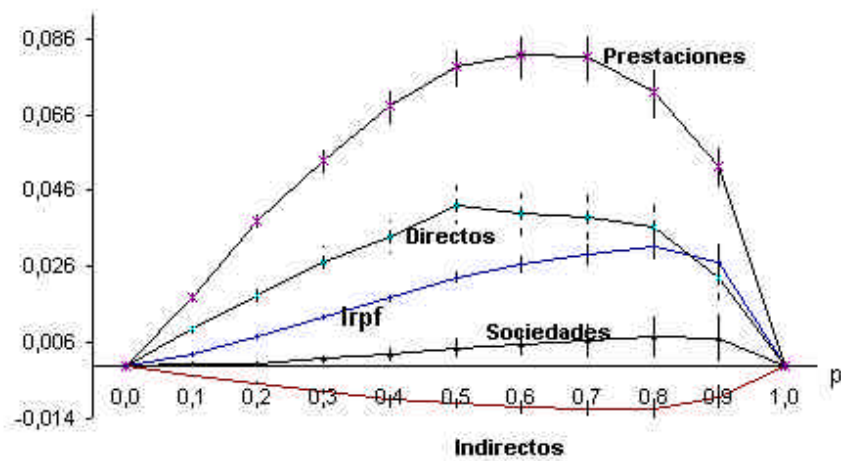


GRÁFICO 4

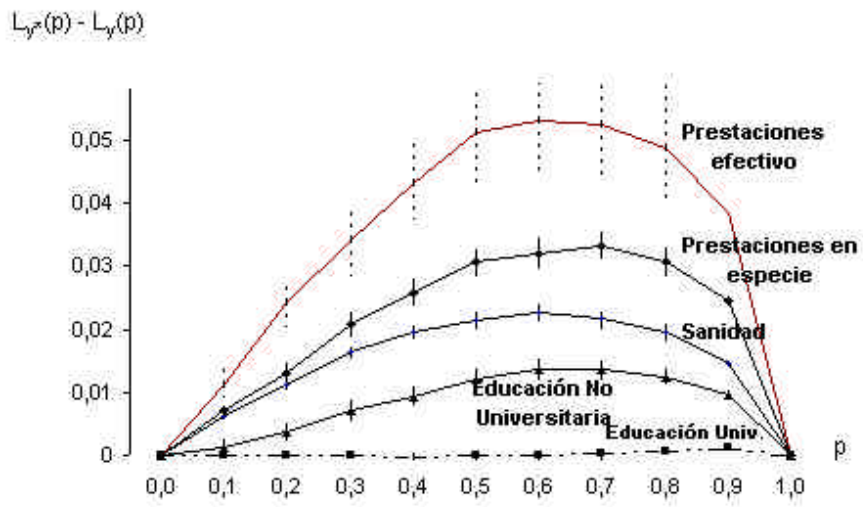


GRÁFICO 3

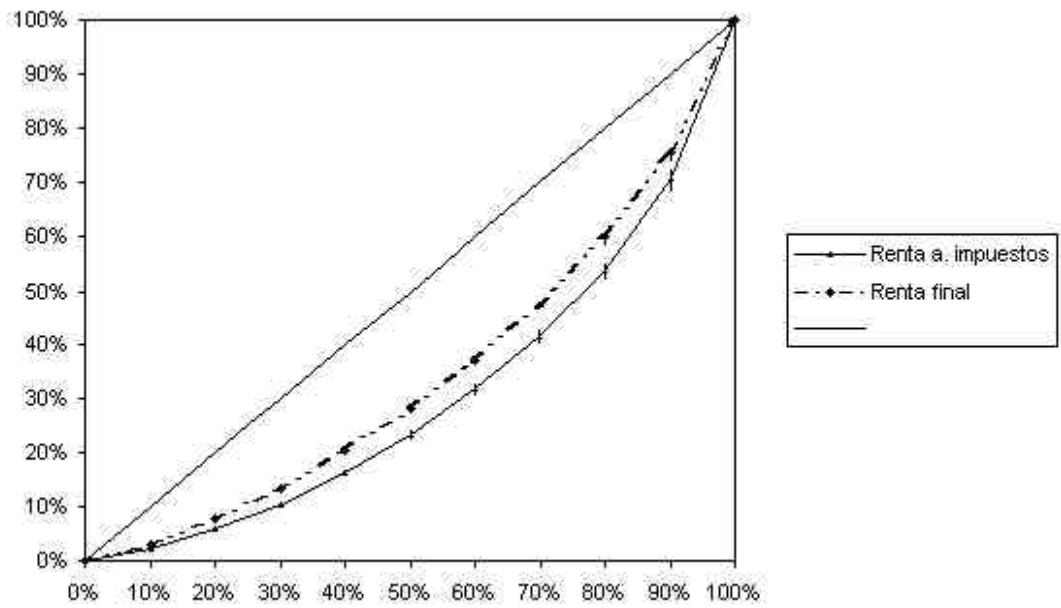


GRÁFICO 5

$$IR(p) - IR(p,-1)$$

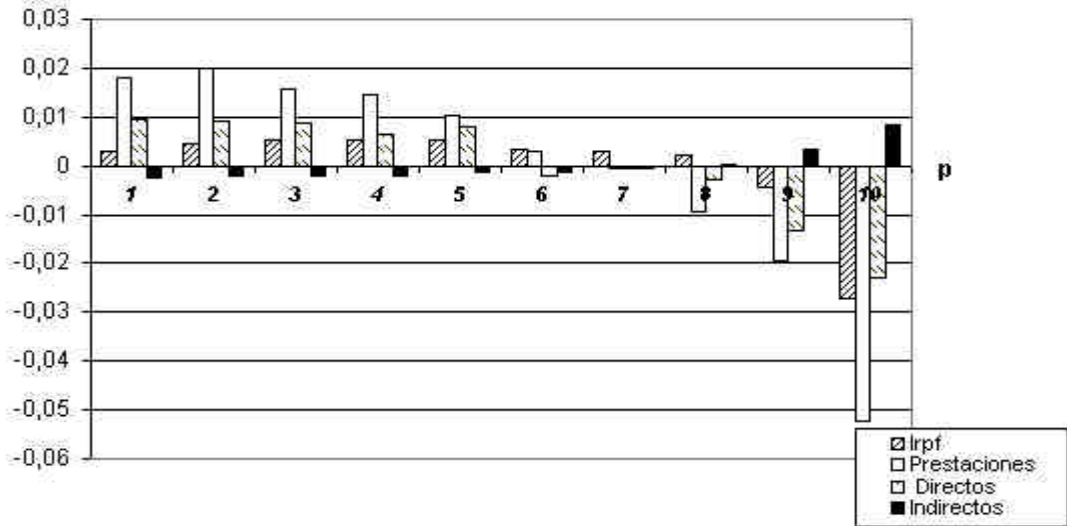
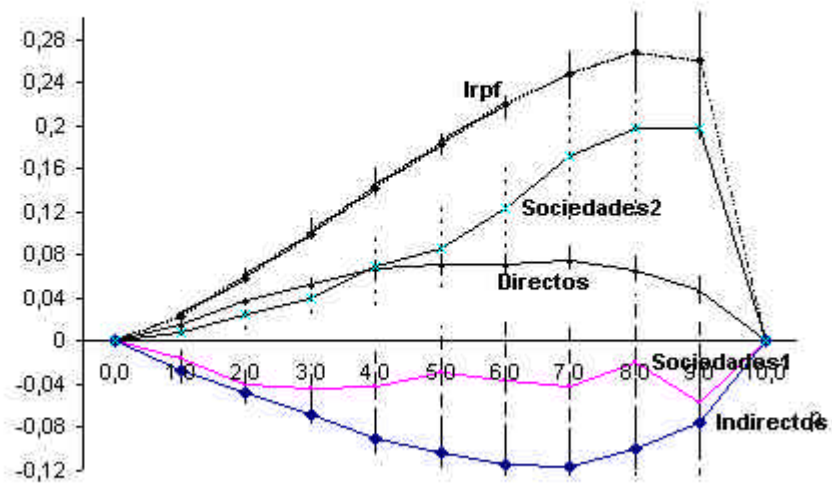


GRÁFICO 6

$$L_y(p) - L_{y,T}(p)$$



Referencias Bibliográficas

- Andres, R (2001) Indices de progresividad y U-estadísticos. Mimeo CREB
- Alegre J., J. Arcarons, S. Calonge y A. Manresa (2001) El método de la fusión estadística: Una aplicación entre le encuesta de presupuestos familiares (EPF90) y una muestra de declarants del impuesto sobre la renta de las personas físicas (IRPF90) en *Desigualdad, Redistribución y Bienestar: Una Aproximación a partir de la Microsimulación de Reformas Fiscales* Labeaga, J.M y M. Mercader (coords.) Instituto de Estudios Fiscales. Madrid.
- Argimón, I y J.M. González Páramo (1987). Traslación e incidencia de las Cotizaciones Sociales por niveles de renta en España, 1980-1984», *Fundación Fondo para La Investigación Económica y Social, D.T. n1.*
- Bandrés, E (1990) Los efectos de los Gastos sociales sobre la distribución de la renta en España, Instituto de Estudios Fiscales, Madrid
- Bishop,J, Formby, J and Zheng, B (1998) Inference Tests for Gini-Based Tax Progressivity Indexes *Journal of Business & Economics Statistics*, Vol 16, No.3, 322-330
- Calonge, S y A. Manresa (2001) *Incidencia Fiscal y del Gasto Público Social sobre la Distribución de la Renta en España y sus CC.AA.* Fundación BBVA. Bilbao.
- Calonge,S y A, Manresa (1997). Consecuencias redistributivas del Estado del Bienestar en España: Un análisis empírico desagregado *Moneda y Crédito*, 204: 13-65.
- Dadvison, R and J-Y, Duclos (1997) Statistical Inference for the Measurement of the incidence of Taxes and Transfers *Econometrica* Vol 65, No. 6 , 1453-1465
- Gimeno, J (1993) Incidencia del Gasto Público por niveles de renta (España 1990 vs 1980), en *Actas del I Simposio sobre Igualdad en la distribución de la Renta y la Riqueza* Fundación Argentaria.
- Estruch, A (1996) Desigualdad y Política Redistributiva. Una estimación de la Incidencia de los Gastos Sociales en España a partir de la Encuesta de Presupuestos Familiares (1990-91). Editorial CEDECS, Economía, Barcelona, 1996
- Lasheras, M, Rabadán, I y R.Salas (1993) Política Redistributiva del IRPF entre 1982 y 1990, en *Actas del I Simposio sobre Igualdad en la distribución de la Renta y la Riqueza* Fundación Argentaria
- Manresa, Calonge y Berenguer (1996). Progresividad y Redistribución de los Impuestos en España . *Papeles de Economía Española* 69:145-159.
- Mayo, R. y R.Salas (1989). Incidencia redistributiva del IVA. Tipos efectivos declarados. (1989) *Hacienda Pública Española* pp. 133-173.
- Mills , J y S, Zandavkali (199/) Statistical Inference via Bootstrapping for Measures of Inequality *Journal of Applied Econometrics*, vol12, 133-150
- Oliver J., J. Ramos, X, y Raymond, J (2001) La mejora en distribución de la renta en España, 1985-1996 en *Desigualdad, Redistribución y Bienestar: Una Aproximación a partir de la Microsimulación de Reformas Fiscales* Labeaga, J.M y M. Mercader (coords.) Instituto de Estudios Fiscales. Madrid.