ENCUESTAS DE PRESUPUESTOS FAMILIARES, RENTA DE LAS FAMILIAS Y ESTUDIO DE LA DISTRIBUCIÓN PERSONAL DE LA RENTA: UNA EXPERIENCIA ESPAÑOLA.

PENA TRAPERO, J.B.
CALLEALTA BARROSO, F.J.
NÚÑEZ VELÁZQUEZ, J.J.
Universidad de Alcalá

1 Introducción.

El trabajo que aquí se presenta es el fruto de un intenso esfuerzo realizado a lo largo de dos años por un equipo de investigación, básicamente de la Universidad de Alcalá de Henares que con el apoyo y colaboración del Instituto Nacional de Estadística y la financiación de la Fundación BBV emprendieron la aventura de estudiar la Distribución Personal de la Renta en España.

Tal vez el término aventura pueda resultar exagerado. Sin embargo, si se tiene en cuenta la limitación de los datos disponibles, se comprenderá mejor el significado que aquí damos a este termino. Aventura es ante todo riesgo y nosotros nos hemos arriesgado a extraer, de la información heterogénea, incompleta y con claras muestras de ocultación, una aproximación a una realidad que quizás pueda ofrecer ciertas dudas, dadas las discutibles hipótesis utilizadas, pero que pensamos será, en todo caso, más cercana a la realidad que la que saldría directamente de las fuentes utilizadas sin un proceso previo de depuración.

Esta comunicación presenta resumidamente parte de los aspectos metodológicos que se tuvieron en cuenta para la realización de dicho estudio. En primer lugar, tras presentar el concepto de renta personal adaptado, se aborda la problemática de decidir sobre la fuente de datos seleccionada y concluir con la necesidad de la corrección de la ocultación detectada en los datos de renta derivados de las Encuestas Básicas de Presupuestos Familiares al compararlos con los deducidos a partir de las Contabilidades Nacionales.

Posteriormente se ha realizado un estudio amplio de los diferentes modelos probabilísticos utilizados en la literatura para modelar las distribuciones de renta, así como de los comportamientos de los principales métodos de estimación y sus consecuencias han sido tenidas en cuenta para su ajuste a los percentiles empíricos de las distribuciones corregidas.

A continuación se hace un detallado estudio teórico del problema de la medición de la desigualdad, de las dominancias de Lorenz y de Lorenz-generalizada y de las múltiples medidas que se ha propuesto para el conocimiento de aquélla. En base a sus consecuencias, se retiene un número limitado de aquéllas que se consideran más características de las diferentes escuelas.

Conscientes de que el análisis de la Distribución Personal de la Renta no se agota con la obtención de las distribuciones y con las medidas de desigualdad, los dos capítulos siguientes abordan aspectos que nosotros consideramos de notable interés, como el estudio de los factores

condicionantes de la distribución, entre los que se han tratado los factores macroeconómicos, la edad, el sexo, el nivel de estudios y las categorías socioprofesionales.

Finalmente se ha intentado hacer un estudio de la movilidad de las rentas que, aunque de naturaleza limitada y en algunos casos puramente metodológica, ha permitido llegar a resultados interesantes.

2 El concepto de renta personal adoptado

La Contabilidad Nacional nos permite diferenciar y matizar el concepto de renta. Así, se debe distinguir claramente entre la renta como retribución de factores, que constituye lo que denominamos **renta primaria**; o la renta final que les queda a las familias o los individuos después de los procesos redistributivos que se desarrollan en las sociedades modernas para corregir o modificar las reglas del mercado y que dan lugar a lo que se llama **renta disponible**. Este último concepto está más ligado a los usos que los individuos o las familias puedan hacer de los ingresos. Permite, además, estudiar los mecanismos de redistribución, elemento básico para enjuiciar el comportamiento de una sociedad en relación con una determinada idea de justicia distributiva. Por ello hemos decidido adoptar en nuestro estudio, como concepto de renta, este último de renta disponible.

Sin embargo, a la hora de concretar la **distribución personal** de la renta, debemos considerar el hecho familiar como redistribuidor de las rentas primarias de los perceptores entre los propios miembros de las familias y tomar las medidas metodológicas necesarias para asignar a cada individuo su renta personal disponible imputada.

Una forma bastante corriente de atacar este problema es utilizar Escalas Equivalentes de Consumo como coeficientes redistribuidores. Nosotros hemos desechado este procedimiento y hemos empleado simplemente el tamaño de las familias como medida de homogeneización, basando nuestra decisión en las siguientes razones:

a) La amplia literatura sobre escalas de consumo todavía no ha llegado a fijar un criterio para determinar una escala que tenga aceptación general. Las líneas metodológicas seguidas son muy diferentes, lo que dificulta la unificación de las escalas o, al menos, una aproximación suficiente para poder recomendar en base científica qué escala hay que utilizar. Con frecuencia se emplea aquélla que es de más fácil aplicación (escala de Oxford, de la OCDE, etc.) pero sin ninguna justificación teórica o empírica.
b) Las dudas suscitadas en el punto anterior son tan ciertas que una de las personas que más han trabajado estos temas -Julie A. NELSON-, después de un estudio empírico sobre los métodos de sistema de las escalas de equivalencia llega a la siguiente conclusión: "A pesar de la elegancia con la que los modelos de BARTEN y los de ROTHBARTH puedan ser expresados teóricamente, es prematuro asumir que desde el punto de vista empírico y práctico, tales modelos generen estimaciones de escalas de equivalencia de más alta calidad que aquellos generados por métodos ad hoc más sencillos" (Income and Wealth, serie 38, nº 3, Sept. 1992, pág. 306). Y en otra parte: "La búsqueda de un verdadero y definitivo conjunto de escalas parece una quimera ya que no existe un modelo completamente superior que podamos estimar" (Household equivalence scales: Theory versus policy Journal of Labor Economics. Vol.11, núm. 3, pág. 489 (1993)).
c) Las escalas de consumo podrían ser aplicables en estudios de demanda, pero hay que tener en cuenta que el consumo no es la renta.
d) El razonamiento subyacente en la utilización de las escalas de equivalencia de consumo en los

estudios de la distribución personal de la renta es la aceptación de la igualdad entre renta y bienestar y entre éste y el consumo. Estos conceptos son, sin embargo, muy diferentes y por ello, de utilizar escalas de equivalencia deberían de ser distintas. Así lo demuestran claramente R.A. POLLAK y F.J. WALES ("Welfare comparaisons and Equivalence Scales", American Economic Review, 1979). Estos autores inician el artículo diciendo: "En este trabajo nosotros defendemos que las escalas de equivalencia requeridas para comparaciones de bienestar son lógicamente distintas de aquéllas que surgen en el análisis de la demanda" y llegan a la conclusión de que dado que el bienestar no tiene carácter aditivo, no se pueden hacer comparaciones de bienestar entre familias de distintos perfiles demográficos.

Por todas estas razones nos ha parecido conveniente trabajar en términos "per cápita", que es un concepto claro, preciso y sencillo. La hipótesis que se hace al utilizar la renta disponible por persona es asumir que la masa de renta disponible por la familia está al servicio de todos los miembros de la misma por un igual sin ponderaciones diferenciadoras según el uso que los miembros puedan hacer de la misma. Se piensa que tal uso no es exclusivamente el consumo, ya que también el ahorro es un componente de la renta y que el nivel de bienestar asociado al uso de la renta no se mide adecuadamente de forma exclusiva por el consumo.

3 *El marco empírico.*

Existen dos posibles fuentes de información acerca de las rentas de las personas físicas: las de origen fiscal y las proporcionadas por las Encuestas a Hogares sobre Presupuestos Familiares. Se han explorado ambas fuentes y aunque en el estudio se hace una referencia a los resultados deducidos de la información fiscal del Impuesto sobre la Renta de las Personal Físicas, el núcleo básico del trabajo se apoya en los datos derivados de las Encuestas de Hogares (1973, 1980 y 1990) que cubren un período de cerca de 20 años, lo que permite una visión temporal suficientemente larga para apreciar los cambios producidos en la distribución de la Renta.

3.1 Las fuentes fiscales.

Al analizar la posibilidad de utilizar las informaciones procedentes de las declaraciones del Impuesto sobre la Renta de las Personas Físicas para determinar la Distribución Personal de la Renta, se encontraron una serie de graves limitaciones que aconsejaron su no consideración. Entre ellas, las siguientes:

a) Los datos de las declaraciones del Impuesto sobre la Renta de las Personas Físicas adolecen de los fuertes sesgos que introducen en los mismos el fraude fiscal existente. En los años de referencia del estudio, esos niveles de fraude, según estimaciones realizadas, se elevaban globalmente por encima del 40% en los rendimientos declarados procedentes del trabajo y eran superiores al 70% en los restantes ingresos declarados, es decir, en la suma de los rendimientos del capital, actividades profesionales, empresariales y agrarias y restantes rendimientos.
b) Al mismo tiempo, se desconoce la distribución de ese fraude en las distintas clases o intervalos de renta, pues esa distribución depende no sólo de conductas individuales difícilmente tipificadas, sino también de la propia estructura de la composición de la renta por tramos, dado que los rendimientos del trabajo son más fácilmente controlados por la Administración Tributaria.

c) Al mismo tiempo, hay que tener en cuenta que algunos de los cambios aparentes en la distribución pueden estar provocados por las propias medidas fiscales de aumento de los mínimos exentos o de la mayor presión de las actividades de control sobre ciertos tramos de renta

Todo ello, unido a los efectos de la inflación difícilmente corregibles sobre una parrilla de clases de intervalos ya dados y de compleja alteración, hace que las conclusiones que podían derivarse de los datos fiscales sobre distribución de la renta declarada fueran escasamente fiables.

3.2 Las Encuestas Básicas de Presupuestos Familiares.

Desde la perspectiva del estudio de la distribución personal de la renta en España, podemos justificar la elección de los datos procedentes de dichas encuestas básicas de presupuestos familiares, basándonos en la riqueza de su información, en lo elemental de sus unidades muestrales (familias, cuya composición es conocida), así como en la fiabilidad de los procedimientos de recogida y contrastación (se recogen a la vez ingresos y gastos, lo que permite un primer control del nivel de contrastación de la información), lo cual los hacen muy propicios a tal fin.

Sin embargo, al analizar la posibilidad de utilizar las informaciones procedentes de las distintas Encuestas a Hogares sobre Presupuestos Familiares para determinar la Distribución Personal de la Renta, se presentaron una serie de limitaciones y cuestiones que cuando fue posible se trataron de resolver razonadamente y otras simplemente se tuvieron en cuenta para no incurrir en abusos que llevaran a conclusiones falseadas.

Entre las limitaciones, es evidente que el estudio de los factores sólo tiene sentido si se refiere al concepto de **rentas primarias** percibidas por el individuo. Y para ello es necesario utilizar Encuestas de Hogares que permitan conocer los ingresos individualizados de cada componente del hogar. Pero de entre las grandes Encuestas Básica de Presupuestos Familiares de 1973, 1980 y 1990, las dos primeras sólo tenían información de los ingresos globales de las familias y solamente la de 1990, daba, además, información individualizada de cada perceptor. Además el hecho de que las Encuestas Básicas de Presupuestos Familiares no sean en cuanto a su diseño encuestas de panel, impiden realizar análisis dinámicos sobre las mismas, pudiéndose tan solo establecer exclusivamente algunos planteamientos de comparativa estática. Es fundamentalmente desde esta óptica desde la que han sido utilizadas en el estudio.

Cuando hemos planteado algún análisis de tipo puramente dinámico, nos hemos visto abocados a utilizar necesariamente las Encuestas Continuas de Presupuestos Familiares, la mejor de las alternativas posibles a este fin, que aunque no están concebidas con esta finalidad y como consecuencia tienen un tamaño muestral reducido para abordar esta problemática con garantías plenas de fiabilidad, al menos nos han permitido una aproximación metodológica ilustrativa desde 1985 a 1993, para el análisis de la evolución temporal del comportamiento de los factores.

En cuanto a la pregunta ¿Por qué utilizar los datos de ingresos y no los del gasto de las Encuestas de Hogares?, existen numerosos trabajos sobre distribución de la renta que se han apoyado en las distribuciones del gasto deducido de las Encuestas de Hogares. Los que así hacen razonan en el sentido de que:

a) Estadísticamente los datos sobre el gasto son más fiables que los datos sobre los ingresos.
b) El gasto puede ser una buena aproximación a la renta permanente.

Frente a estos argumentos que son parcialmente válidos, nosotros pensamos que la renta y el gasto son conceptos relacionados pero obviamente muy diferentes y entre nuestros argumentos están los siguientes:

a) Aproximar la renta a través del gasto sólo es aceptable para los niveles bajos de renta en que la propensión marginal al consumo es próxima a uno. A medida que nos desplazamos a los niveles altos de la renta, ésta se separa cada vez más del gasto y la distribución de ambos agregados se hacen cada vez más diferentes. Baste pensar que mientras el gasto tiene un límite marcado por el nivel de saturación, la renta tiene un límite teóricamente infinito.
b) La utilización de la distribución del gasto como aproximación a la distribución de la renta introduce un sesgo importante en los tramos altos y desvirtúa el análisis. Si se hace pensando que es una forma de corregir la ocultación de los ingresos creemos que el método de corrección es inadecuado, impreciso y falso.
c) Tampoco nos convence el argumento de que el gasto aproxima bien el concepto de renta permanente porque la renta transitoria también es renta y no se ve el porqué hay que excluirla.
d) Otro argumento que va en la misma dirección que lo señalado la proporciona N. KAKWANI (Income Inequality and Poverty -1980- pág. 167 y ss.) que demuestra que "el gasto de consumo personal es superior Lorenz al ingreso disponible". Esto implica que el consumo personal se distribuye de forma más igualitaria que la renta disponible.
e) Finalmente, conviene señalar que dado el diseño de las Encuestas de Consumo realizadas por el I.N.E., los consumos familiares analizados no son los consumos reales de las familias, debido a que los gastos tienen distintos períodos de referencia y se les aplica un factor de elevación diferente para analizarlos. Así por ejemplo, si una familia declara que ha comprado un bien de periodicidad mensual - pongamos un abrigo- se multiplica por 12 ese gasto al analizar el consumo de una familia en compensación de aquellas familias que han sido entrevistadas en el verano y que no compran lógicamente vestidos de invierno. Se trata en definitiva de un diseño muestral bidimensional referido al tiempo y al espacio.

En cuanto a la homogeneidad de conceptos a lo largo de los años del estudio, existe continuidad metodológica para las encuestas Continuas y una compatibilidad aceptable en las encuestas Básicas. Se ha estudiado la compatibilidad de definiciones y clasificaciones relativas a los hogares, las personas, los ingresos y los gastos en las EBPF. No obstante, en ocasiones ha sido necesario elaborar clasificaciones compatibles a lo largo del tiempo a partir de las definiciones de las encuestas de cada año, especialmente para las clasificaciones en categorías socioeconómicas

3.3 La Cuenta de Rentas del Sector Hogares.

La renta disponible es el saldo de la cuenta de renta de las familias, por eso nuestra metodología nos ha llevado a tratar de encuadrar esa magnitud dentro del marco conceptual que fija la Contabilidad Nacional y, con esta finalidad, se ha realizado un laborioso y meticuloso trabajo que podríamos denominar de enlace micro-macro.

Después de un minucioso proceso de análisis de los conceptos de las tres Encuestas Básicas y de los conceptos análogos de la cuenta de renta de las familias de la Contabilidad Nacional, se procedió a establecer la correspondencia aproximada entre los conceptos que integran la Renta Bruta Disponible de los hogares según la Contabilidad Nacional y las Encuestas Básicas de

Presupuestos Familiares. Esta correspondencia se pudo establecer mediante dos reglas de homogeneización diferentes: una para las encuestas de 1973 y 1980 y otra para la más moderna de 1990.

A partir de estas informaciones teóricamente comparables, tratamos de describir y analizar la distribución personal de la renta disponible, cuya forma probabilística se derivara de la que presentaban las Encuestas de Hogares en la parte de los ingresos y cuya renta agregada fuera compatible con la derivada de la cuenta de hogares de la Contabilidad Nacional.

3.4 Corrección de la ocultación.

Se constatan grandes discrepancias entre las Rentas deducidas a partir de los datos procedentes de las Encuestas Básicas de Presupuestos Familiares y los datos de referencia deducidos de las Contabilidades Nacionales, a la vez que se observan deficiencias importantes en aquellas mismas rentas deducidas que hacen pensar sin duda en un alto grado de ocultación en las profesiones liberales y autónomas, así como en las rentas de la propiedad

La comparación de la renta disponible deducida de las Encuestas y la que para los mismos años señalaba la Contabilidad Nacional indicaba una infraestimación por parte de las Encuestas del orden del 40%. Por otra parte, parecía razonable suponer que el grado de subdeclaración no afectaba únicamente al nivel de la renta, sino que afectaba también a su forma, ya que la subdeclaración era muy diferente según los distintos tramos. Esto nos llevó a plantearnos la necesidad de proceder a una corrección de los datos originales, a fin de obtener otra distribución que, respetando lo más posible su forma y características, fuera congruente con los agregados que finalmente han sido deducidos de las Contabilidades Nacionales para la renta total repartida en España, cada comunidad autónoma, cada categoría socioprofesional o cada clase de hábitat considerado en dicho estudio anterior.

Basar el estudio en la información original sin correcciones por subdeclaración, nos llevaría a reflejar una realidad falsa, por muy buenos que fuesen los métodos de análisis. Falsa en los niveles y falsa en la distribución. Hemos, pues, preferido arriesgarnos a efectuar correcciones basándonos en hipótesis razonables, aunque lógicamente discutibles.

Las correcciones se han efectuado en la doble dirección de obtener un nivel adecuado según las diferentes desagregaciones por ámbitos regionales, socioprofesionales y clase de hábitat, y de lograr, una vez fijados los niveles, un reparto aceptable entre los individuos. Básicamente el proceso de corrección ha sido el siguiente:

a) Respecto a los niveles: Partiendo de la propia información de la Encuesta sobre gastos, ingresos y ahorro, se obtuvieron coeficientes de ocultación para cada desagregación realizada. Aplicados dichos coeficientes a los ingresos declarados se obtuvo una masa de ocultación que se repartió entre las distintas partidas de la cuenta de renta de las familias en función de distintos criterios basados en información adicional o, si no se disponía de ella, del nivel de la partida antes de la corrección. De esta forma se obtuvo una cuenta de renta de las familias que sería la que podríamos llamar "óptima" en el sentido de ser la mejor que se podría obtener a partir de las Encuestas. Esta cuenta de renta se consideró una estructura "óptima" desagregada por Autonomías, categorías socioprofesionales y por clase de hábitat.

Una vez fijada esta estructura se procedió a repartir, proporcionalmente partida a partida, la cuenta

de renta de las familias que da la Contabilidad Nacional para los años 1980 y 1990. Para el año 1973 ha sido preciso realizar un estudio previo para separar la cuenta de renta de las familias que estaba unida a la cuenta de renta del sector institucional S.10 -Empresas no financieras-.

b) Respecto a la forma de la distribución: Una vez determinados los niveles homologables con los de la Contabilidad Nacional, la diferencia entre aquéllos y los que daban las Encuestas, por cada desagregación, determinó una masa de Renta oculta que había que repartir entre los individuos.

Como una medida simple, hemos considerado, para abordar el problema de corregir nuestros datos, la hipótesis de que la ocultación debe recaer más en las clases de mayores ingresos, ya que serán estas las que podrán acceder más fácilmente a la propiedad de bienes; y por tanto, a mayor nivel de ingresos, mayor nivel de ocultación debemos imputarle. Por ello hemos preferido utilizar el procedimiento más simple que con la información disponible puede ofrecernos una aproximación global válida a la imputación de ingresos ocultados a las familias. Hemos realizado esta imputación basándonos en la hipótesis inicial que asegura que las cantidades ocultadas en cada clase son proporcionales a la renta repartida en la misma.

Sin embargo, y puesto que disponemos de las cantidades ocultadas en cada región, en cada categoría socioprofesional y en cada clase de hábitat, la corrección se hará respetando la pertenencia de las familias a estas clases, y teniendo en cuenta la ocultación global que en ellas se dio.

En primer lugar se ha tratado de establecer una única corrección de la distribución conjunta para todos los individuos que respetase todas las correcciones necesarias en cada región, categoría socioprofesional o clase de hábitat. Sin embargo, ninguno de los siete intentos de asignación de las cantidades corregidas a nivel de microdato que hemos estudiado con este objeto ha tenido éxito. Y dado que los estudios que debíamos derivar a partir de entonces iban a centrarse sobre las clasificaciones marginales según la región, categoría socioprofesional, clase de hábitat o sobre el conjunto nacional agregado, hemos decidido finalmente realizar tres correcciones distintas de cada dato en función de su clase de pertenencia, en cada uno de los tres tipos de distribuciones marginales (región, categoría socioprofesional y clase de hábitat), y calcular una cuarta corrección para la distribución agregada a nivel nacional, como el promedio de las correcciones realizadas sobre cada dato por pertenecer a su región, su categoría socioprofesional y su clase de hábitat. Estas nuevas correcciones así planteadas deberían respetar los totales ocultados por clase y el total ocultado nacional, así como también respetan el principio de proporcionalidad a los ingresos declarados (a más ingresos, más ocultación).

Hemos planteado, pues, una hipótesis de progresividad lineal en la propensión a la ocultación a medida que los ingresos aumentan, de modo que cuando estos aumentan, no sólo aumente la cantidad ocultada, sino también la intensidad o proporción de la cantidad ocultada. Así hemos planteado el modelo general de *Corrección con Tasa de Ocultación lineal progresiva.* Sin embargo, hemos adoptado como hipótesis final de trabajo una situación intermedia entre las dos situaciones extremas entre los casos de mínima progresividad (tasa de ocultación constante) y máxima progresividad lineal en la propensión a la ocultación, compatibles con las ocultaciones estimadas.

Así, en el caso de admitir como pendiente m de la recta que explique la tasa de ocultación, una fracción prefijada 1/k de la máxima pendiente posible,

$$m = \frac{1}{k} \frac{O}{\sum_{clase} x^2},$$

obtenemos que:

$$O_x = x \bullet \frac{O}{k} \left\{ \frac{k\text{-}1}{\sum_{clase} x} + \frac{x}{\sum_{clase} x^2} \right\} \quad y \quad X = x \left(1 + \frac{O}{k} \left\{ \frac{k\text{-}1}{\sum_{clase} x} + \frac{x}{\sum_{clase} x^2} \right\} \right)$$

siendo O_x la ocultación imputada a una renta x, cuando en su *clase* se ocultó O, y X la renta finalmente corregida.

En la tabla del Anexo 1 adjunto, a modo de ejemplo, bajo la columna *Análisis Empírico*, presentamos los resultados obtenidos con los datos corregidos a nivel nacional, mediante este método, el cual respeta las hipótesis planteadas, además de prever una tasa de ocultación no nula en las rentas inferiores. Todo el análisis posterior de resultados sobre estas encuestas básicas se han derivado de estas distribuciones corregidas en el sentido arriba indicado.

4 *La modelización probabilística de distribuciones de ingresos.*

La regularidad mostrada por las distribuciones de ingresos observadas en el tiempo y en el espacio, justifica suficientemente el intento de describirlas con ayuda de alguna función de distribución estadística. Esto provee no sólo un procedimiento de resumir útilmente el fenómeno de la distribución de ingresos, sino también una técnica para estudiar los efectos de alternativas políticas redistributivas: proporciona un instrumento para interpolar dentro de cada clase de ingresos, y para extrapolar en clases más bajas y más altas; un procedimiento para obtener los indicadores de desigualdad a partir de los parámetros estimados, sin la necesidad de admitir igualdad de ingresos dentro de las clases; el establecimiento de relaciones existentes entre distintas medidas de desigualdad, de valiosa ayuda para interpretar dichas medidas; la posibilidad de ser utilizada para simular un proceso complejo en el que los efectos a observar dependan de las sumas de los efectos personales de las unidades económicas integrantes; la posibilidad de construcción de modelos cuyo objetivo sea predecir variables específicas, que dependan del conjunto de la distribución de ingresos a través de la investigación de estos parámetros solamente; etc.

Se observa que las distribuciones de ingresos son siempre asimétricas positivamente, con una moda simple, y con una larga cola contactando débilmente con el eje de abscisas, siendo muchos los trabajos que muy diversos autores han desarrollado sobre esta hipótesis, entre los que destacan los de Pareto, Kapteyn, Edgeworth, Gibrat, Champernowne, Salem y Mount, Singh y Maddala, McDonald, etc. hasta llegar a los más recientes de Colombi, Stoppa, Dagum y otros. En todos ello hay siempre un denominador común: el interés por encontrar la distribución que mejor represente la forma general de los datos nacionales o regionales considerados. Pero, lamentablemente, en muy pocas ocasiones los datos considerados en aquéllos son los de nuestro país, y muchas menos, los de nuestras regiones; y cuando ello ocurre, no se llega más allá que al, no por ello menos meritorio, intento de ajuste a algún tipo de distribución básica predeterminada.

A continuación nos planteamos un triple objetivo. Por un lado, examinar estadísticamente las

distintas distribuciones que a lo largo del tiempo han sido utilizadas como alternativas para el ajuste de las distribuciones de ingreso. Por otro lado, seleccionar razonadamente los principales modelos probabilísticos candidatos a representar correctamente la distribución del tamaño de los ingresos en España, en base a los datos proporcionados por las Encuestas Básicas de Presupuestos Familiares. Y en tercer lugar, establecer procedimientos adecuados para el ajuste de tales modelos, utilizando los conocimientos anteriormente desarrollados.

4.1 Los modelos teóricos de las distribuciones de ingresos.

Hemos organizado las distribuciones que generalmente se han utilizado para este fin (excluidas mixturas), en 4 familias que faciliten el estudio de sus propiedades, además de sus posibles transformadas.

Familia de Pearson: Pearson en 1948 plantea su sistema de distribuciones, a partir de la resolución de la ecuación diferencial siguiente, siendo f una función de densidad definida sobre [u,v) Real.

$$\frac{df}{dx} = \frac{(x-a)\cdot f(x)}{(b_0 + b_1 \cdot x + b_2 \cdot x^2)}$$

Dependiendo de las raíces del denominador, la ecuación diferencial tiene distintas formas de solución. Por ello, estas formas dependerán de su discriminante, o equivalentemente de la cantidad K=b1**2/(4*b0*b2).

Familia de Perk: Esta familia se caracteriza por tener una función de densidad de la forma:

$$f_y(y) = \frac{\sum_{j=0}^{m} a_j e^{-jqy}}{\sum_{j=0}^{m'} b_j e^{-jqy}} ; a_j , b_j \; y \; q \; reales$$

Familia de D'Addario: Se generan a partir de una variable aleatoria X con función de densidad:

$$g(x) = \frac{A}{b + e^{x^{\frac{1}{p}}}} , x > 0 , p > 0 , A > 0 , b > \_1$$

mediante la transformación $Y=h^{-1}(X)$, solución de la ecuación diferencial:

$$x^q . \frac{dx}{dy} = \frac{a}{y - c} , c \le y_0 \le y < +\infty , a \ne 0 , q \in \Re$$

La función de densidad g(x) es monótona decreciente, y la solución de la ecuación diferencial $Y=h^{-1}(X)$, es monótona creciente o decreciente dependiendo del signo de α, por lo que la composición de las dos abarca un amplio rango de distribuciones unimodales, en su origen o en el interior de su rango.

Familia de distribuciones que buscan el cumplimiento de características regulares observadas

empíricamente en las distribuciones de ingresos: Si la ley Fuerte de Pareto dice que la elasticidad r(x) de la función 1-F(x) debe ser constante, la ley débil exige sólo que esto se cumpla en el límite, cuando x tienda a infinito. Así pues, una posible vía para encontrar alguna de tales distribuciones será la resolución de la ecuación diferencial:

$$r(x) = \frac{x.\frac{dF}{dx}}{1 - F(x)} = \frac{a \cdot x^b}{c^b + x^b}, con\ a > 0, b > 0, c > 0$$

cuya solución, los momentos de orden j con respecto del origen se demuestra que sólo existen cuando $j<\alpha$. Empíricamente, se observa que α toma valores entre 1 y 3 generalmente, lo que preserva la existencia de 1 ó 2 momentos finitos.

Sin embargo, estas familias pueden ser generadas a partir de un sistema de hipótesis más amplio, propuesto por Dagum en 1980, 1983 y 1985, consistentes en: a) La elasticidad ε de la F. de distribución, es decreciente y cóncava de F(y); b) Comienza con un valor finito y positivo cuando $f(y)\rightarrow 0$; c) Y decrece hasta cero cuando $f(y) \rightarrow 1$, al tender y a infinito; pudiéndose obtener pues las distribuciones de esta familia como las soluciones de la ecuación diferencial siguiente, en la que las restricciones aseguran que la elasticidad de F es una función de F (y por tanto, de y) positiva, decreciente y acotada, pudiendo usarse incluso cuando hay ingresos nulos y negativos (considerados nulos) considerablemente significativos. En este caso, $y_0=0$, y $0<\alpha<1$. Cuando la población posea ingresos superiores a una cantidad mínima y_0, entonces, $\alpha<0$ y $F(y_0)=0$.

$$e(F(y) - a, y) = \frac{dlog[F(y) - a]}{dlog(y)} = y(y).f(F) \leq k$$

$$\text{siendo } k > 0, 0 \leq y_0 < y < +\infty, a < 1, y(y) > 0, f(F) > 0$$

$$\text{y donde } \frac{dy(y).f(F)}{dF} < 0, \frac{dy(y)f(F)}{dy} < 0$$

Transformación de variables: Otra vía de obtener distribuciones asimétricas candidatas a la representación de las distribuciones de ingreso, consiste en el empleo de transformaciones monótonas de estas variables aleatorias, para convertirlas en variables simétricas, y más concretamente normales. Siguiendo a Johnson y Kotz, dada una distribución X, dependiendo del tipo de rango de la distribución, podemos intentar distintas transformaciones de la misma para hacer que se aproxime a una distribución conocida, estudiando estos autores las propiedades de las mismas, cuando Z es una N(0,1). Así,

- Si su rango se encuentra acotado inferiormente por α, podemos intentar

  $Z=\gamma + \delta \log(X-\alpha)$, $X>=\alpha$, $\delta>0$

- Si su rango se encuentra acotado inferiormente por α y superiormente por $\alpha+\lambda$, siendo $\lambda>0$, podemos intentar

  $Z=\gamma + \delta \log((X-\alpha)/(\alpha+\lambda-X))$, $\alpha<=X<=\alpha+\lambda$, $\delta>0$

- Si su rango no se encuentra acotado, podemos intentar

  $Z=\gamma + \delta \sinh^{-1}((X-\alpha)/\lambda)$

Como conclusión a nuestro primer objetivo y como base para cubrir el segundo, hemos identificado más de 30 distribuciones que por una u otra vía han sido utilizadas para la modelización de las distribuciones de ingresos, para cuyo estudio hemos tenido que estudiar detalladamente algunas más que servían de pivote en su generación. Para cada una de ellas se ha confeccionado una ficha técnica con sus principales características, así como un estudio gráfico del efecto de sus parámetros, con el fin de que nos sirviera de ayuda a la hora de seleccionar morfológica y estadísticamente las candidatas más prometedoras de entre ellas.

4.2 Consideraciones teóricas para la Selección de un modelo

Un primer paso obvio en la reducción del conjunto de distribuciones estudiado ha sido el identificar aquellas variables que pueden considerarse como casos particulares de otras más generales. En este sentido, y en las siguientes tablas recogemos las dependencias entre las mismas.

Familia de Pearson y sus Transformadas.

PEARSON TIPO I	$\rightarrow$ BETA GENERALIZADA (p,q;a,b)
PEARSON TIPO VI	$\rightarrow$ F-SNEDECORD (η_1,η_2)
PEARSON TIPO IV	$\rightarrow$ LOG-PEARSON IV ($\alpha_1,\alpha_2,\alpha_3,\alpha_4$,me)
NORMAL (μ,σ)	$\rightarrow$ LOG-NORMAL ($\mu,\sigma;\theta$)
NORMAL (0,1)	$\rightarrow$ N.S.H.I ($\beta_1,\beta_2,\beta_3,\beta_4$)
PEARSON TIPO VII	$\rightarrow$ LOG-STUDENT (α,η,me)

Familia de Perk y sus Transformadas.

LOGÍSTICA (μ,σ)	$\rightarrow$ LOG-LOGÍSTICA (μ,σ)
CHAMP (α,λ,μe)	$\rightarrow$ CHAMPERNOWNE (α,θ,me)
CHAMP (α,λ,μe)	$\rightarrow$ BOX-COX-CHAMPERNOWNE (α,θ,me,β)

Familia de D'Addario.

AMOROSO 4P ($p,\lambda,s;y_0$)	
GAMMAG 4P ($\alpha,\beta,c;y_0$)	= AMOROSO 4P ($\alpha,\beta^{-c},1/c;y_0$)
AMOROSO 3P (p,λ,s)	= AMOROSO 4P ($p,\lambda,s;0$)
GAMMA 3P ($\alpha,\beta;y_0$)	= AMOROSO 4P ($\alpha,\beta^{-1},1;y_0$)
WEIBULL ($\lambda,s;y_0$)	= AMOROSO 4P ($1,\lambda,s;y_0$)
VINCI (p,λ)	= AMOROSO 4P ($p,\lambda,1;0$)
GAMMA 2P (α,β) [March]	= AMOROSO 4P ($\alpha,\beta^{-1},1;0$)
EXP.NEG. ($\beta;y_0$)	= AMOROSO 4P ($1,\beta^{-1},1;y_0$)
GAMMA 1P (α)	= AMOROSO 4P ($\alpha,1,1;0$)
χ^2 (η)	= AMOROSO 4P ($\eta/2,1/2,1;0$)
χ (η)	= AMOROSO 4P ($\eta/2,1/2,1/2;0$)
NORMAL (μ,σ)-mitad derecha	= AMOROSO 4P ($1/2,1/2\sigma^2,1/2;\mu$)
RAYLEIGH (α)	= AMOROSO 4P ($1,\alpha,1/2;0$)
MAXWELL (α)	= AMOROSO 4P ($3/2,\alpha/2,1/2;0$)

Familia de Dagum.

PARETO III (α,c,β;y_0)	
PARETO II (α,c;y_0)	= PARETO III (α,c,0;y_0)
PARETO I (α;y_0)	= PARETO III (α,0,0;y_0)
SINGH-MADDALA (α,β,λ)	
BENINI (α;y_0)	
LOG-GOMPERTZ (α,λ)	
DAGUM II/III (α,λ,δ,β)	
DAGUM I (λ,δ,β)	= DAGUM II/III (0,λ,δ,β)

El criterio "K" de Pearson nos ayuda a detectar incompatibilidades entre las características paramétricas de la distribución empírica y las hipótesis sobre las que se generan las distribuciones teóricas de ingresos de la Familia de Pearson. Así, cabe esperar que para las siguientes distribuciones, el valor de k sea: Beta I (K<0), Normal (K=0), Pearson IV (0<K<1), Vinci (K=1), Beta II y Pareto (K>1), Gamma y exponencial (K tiende a +infinito ó -infinito); de forma que valores distintos descartarían las correspondientes distribuciones.

Desde otro punto de vista, la elección final de un modelo probabilístico particular para la distribución del tamaño de los ingresos debe apoyarse sobre su capacidad de cumplir un cierto conjunto de propiedades económicas, econométricas, estocásticas y matemáticas bien definidas. Así, Aitchison y Brown en 1957, y Metcalf en 1972 ya establecen una lista de cuatro propiedades; Dagum y Bartels separadamente las revisan y amplían en 1977.

Resumidamente, las principales propiedades exigidas serían:

Propiedades acerca del Rango de la Distribución:

1. Acepta un Rango (0,+infinito)
2. Acepta rentas nulas
3. Acepta rentas nulas con probabilidad no nula
4. Acepta rentas negativas
5. Acepta un mínimo no predeterminado sin necesidad de truncamiento

Propiedades acerca de la Forma de la Distribución:

6. Admite Asimetría positiva
7. Puede representar poblaciones unimodales en el mínimo de su rango
8. Puede representar poblaciones unimodales en el interior de su rango
9. Acusa cambios de estas formas a través de los cambios paramétricos

Propiedades acerca de sus Parámetros:

10. Numero de Parámetros (Parsimonia)
11. Interpretabilidad económica de los parámetros (nº de parám.interpretables)

Propiedades acerca de su Fundamentación Económica:

12. Admite la existencia de un número finito de momentos finitos
13. Admite la Convergencia estocástica a la Ley de Pareto
14. Fundamentada en postulados económicos lógico-empíricos

Propiedades acerca de su Manejabilidad Analítica:

15. Facilidad de Manejo analítico (1:Fácil - 9-Difícil)

Por tanto, un segundo paso en la selección de las distribuciones más adecuadas, ha sido, pues, el estudio de estas propiedades para las quince distribuciones esencialmente distintas resultantes de la reducción anterior, y de cuyos resultados se deduce, que el incumplimiento por parte de aquéllas de propiedades como las 1, 6 ó 8, las excluirían automáticamente de ser consideradas, mientras que propiedades como las 9, 11, 12, 13 y 14 son muy deseables y a la vez discriminatorias.

Como consecuencia del análisis realizado de las propiedades de cada modelo, ofrecen unas mejores perspectivas de adecuación para representar las distribuciones de los tamaños de renta, las distribuciones de Champernowne, Singh-Maddala, Log-Gompertz, y Dagum I, II o III, seguidas en un segundo plano por las Gamma Generalizada Triparamétrica, F-Snedecor, Amoroso Tetraparamétrica, Vinci y Weibull, todas ellas con buenas propiedades de carácter empírico-económicas, seguidas finalmente de las Log- Normal, Gamma Generalizada Biparamétrica, Beta Generalizada, Benini, Log-Logística y Log-Student, que si bien no cumplen estos postulados de carácter empírico-económicos o de interpretabilidad de sus parámetros, parecen cumplir bien las demás propiedades de modelado ad-hoc. Teniendo en cuenta las relaciones de inclusión y convergencia estudiadas entre estas distribuciones, cabría concluir con que las distribuciones que deben ser consideradas finalmente para el ajuste de estas distribuciones, son por orden de idoneidad, Champernowne, Singh-Maddala, log-Gompertz, y Dagum II y Dagum III, seguidas en segundo lugar por F-Snedecor y Amoroso tetraparamétrica, y finalmente por log-normal, beta generalizada, Benini y log-Student, si bien para estas últimas, dada su poca fundamentación e interpretabilidad en términos económicos, no existirían argumentos que justificasen el que un posible buen ajuste en un momento dado, tuviese que implicar en ningún sentido que también proporcionara otro buen ajuste en otro instante posterior.

En el siguiente cuadro resumimos el análisis realizado para las propiedades de las distribuciones en estudio:

Tabla 2. Resumen de propiedades de los Modelos Teóricos de la distribución de la renta

Propiedades:	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Familia de Pearson y sus transformadas															
LOG-NORMAL	S	N	N	S	S	S	N	S	N	3	3	N	N	S	2
NORMAL SENHIP INVER	S	S	N	S	N	S	S	S	S	4	N	N	N	N	7
LOG-STUDENT	S	N	N	N	N	S	S	S	S	3	1	N	D	N	5
GAMMAG 3P	S	S	N	S	S	S	S	S	S	4	2	S	D	N	4
GAMMAG 2P	S	S	N	S	S	S	S	S	S	3	2	N	N	N	3
BETA GENERALIZADA	N	S	N	S	S	S	S	S	S	4	3	N	N	N	3
F SNEDECOR	S	S	N	N	N	S	S	S	S	2	2	S	D	N	5
LOG-PEARSON IV	S	N	N	N	N	S	S	S	S	5	1	?	D	N	9
Familia de Perk y sus transformadas															
CHAMPERNOWNE	S	N	N	N	N	S	S	S	S	3	3	S	D	S	6
BOX-COX-CHAMPERNOWNE	S	N	N	N	N	S	S	S	S	4	1	?	N	N	8
LOG-LOGISTICA	S	N	N	N	N	S	S	S	S	2	2	N	D	N	3
Familia de D'Addario															
AMOROSO 4P	S	S	N	S	S	S	S	S	S	4	2	S	D	N	4
AMOROSO 3P	S	S	N	N	N	S	S	S	S	3	1	S	D	N	4
MARCH	S	S	N	N	N	S	S	S	S	2	1	N	N	N	2
VINCI	S	N	N	N	N	S	N	S	N	2	1	S	D	N	2
WEIBULL	S	S	N	S	S	S	S	S	S	3	2	S	D	N	4
GAMMA ESTANDAR	S	S	N	N	N	S	S	S	S	1	1	N	N	N	1

EXPONENCIAL NEGAT	S	S	N	S	S	S	S	N	N	2	2	N	N	N	1
χ^2	S	N	N	N	N	S	S	S	S	1	1	N	N	N	2
χ	S	N	N	N	N	S	S	S	S	1	1	N	N	N	3
NORMAL MITAD DCHA	S	S	N	N	N	S	S	N	N	2	2	N	N	N	5
SEMINORMAL	S	S	N	N	N	S	S	N	N	1	1	N	N	N	5
RAYLEIGH	S	S	N	N	N	S	N	S	N	1	1	N	N	N	3
MAXWELL	S	S	N	N	N	S	N	S	N	1	1	N	N	N	3
Familia de Dagum															
PARETO I	S	N	N	N	S	S	S	N	N	2	2	S	F	S	3
PARETO II	S	S	N	S	S	S	S	N	N	3	2	S	D	S	4
PARETO III	S	S	N	S	S	S	S	N	N	4	1	S	N	S	6
BENINI	S	N	N	N	S	S	N	S	N	2	2	N	N	S	6
SINGH-MADDALA	S	S	N	N	N	S	S	S	S	3	3	S	D	S	5
LOG-GOMPERTZ	S	N	N	N	N	S	N	S	S	2	1	S	D	S	6
DAGUM I	S	N	N	N	N	S	S	S	S	3	3	S	D	S	5
DAGUM II	S	S	S	N	N	S	S	S	S	4	4	S	D	S	5
DAGUM III	S	N	N	N	S	S	S	S	S	4	4	S	D	S	6

4.3 Los procedimientos de ajuste.

Con el objeto de que la selección final del método de ajuste empleado, así como del modelo probabilístico adoptado para ello, no sólo sea lo más fundamentada posible desde un punto de vista teórico, sino también lo más estable posible desde un punto de vista empírico, hemos procedido en primer lugar a estudiar exhaustivamente para el caso nacional, los posibles procedimientos de ajuste de los distintos modelos probabilísticos que hemos considerado, a los datos de rentas per cápita en sus tres versiones estudiadas en los apartados anteriores de este capítulo; a saber, datos básicos procedentes directamente de las Encuestas Básicas de Presupuestos Familiares, los datos corregidos por motivos de ocultación mediante tasa de ocultación constante, y los datos corregidos por motivos de ocultación mediante tasa de ocultación lineal progresiva.

Dada la amplitud de las muestras emplea en las E.B.P.F., se ha optado por la división del recorrido de la variable de ingresos en las 100 clases correspondientes a los intervalos interpercentílicos. Además, hemos empleado como herramienta un programa que proporciona estimaciones paramétricas de mínimos cuadrados (ponderados o no) en modelos no lineales, con restricciones de acotación sobre los parámetros, por lo que ha sido necesario reformular los distintos métodos de estimación que hemos empleado, para hacerlos de factible resolución mediante dicho programa.

Notando F (Frecuencias relativas de las clases), FAC (Frecuencias relativas acumuladas de las clases), DF (Valores de las densidades de frecuencias relativas en las marcas de las clases), FY (Valores de la función de distribución de la variable en el límite superior de la clase), DY (Valores de la función de densidad de la variable en las marcas de las clases), PY (Probabilidades de las clases de división del recorrido de la variable), una vez reformulados, los procedimientos considerados han sido los siguientes:

Método de Máxima Verosimilitud.

Obtener los parámetros θ del espacio paramétrico ω tales que hagan mínima la expresión:

$$\sum_{i=1}^{nc}\left(\sqrt{-F_i.\log(PY_i)}\right)^2$$

Ajuste de Mínimos cuadrados a la Función de Distribución.
Obtener los parámetros θ del espacio paramétrico ω tales que hagan mínima la expresión:

$$\sum_{i=1}^{nc}(FAC_i - FY_i)^2$$

Dagum propone este método como solución a la estimación de máxima verosimilitud de θ cuando consideramos el modelo FAC=FY+ε , estando los residuos distribuidos independientemente según una N(0,σ). Pero esto aquí es poco probable, pues al ser acumulativos tanto FAC como FY, los residuos presentarán el conocido efecto de rachas. Por ello, este método puede ser utilizado por su sencillez para la obtención de una estimación inicial para otro método más eficiente, pero no como método de estimación definitiva en sí.

Ajuste de Mínimos cuadrados a la Función de Densidad.
Obtener los parámetros θ del espacio paramétrico ω tales que hagan mínima la expresión:

$$\sum_{i=1}^{nc}(DF_i - DY_i)^2$$

Cuya solución puede comprobarse que es la estimación de máxima verosimilitud de θ cuando consideramos el modelo DF=DY+ε , siendo ε una variable residual distribuida para cada observación independientemente según una N(0,σ). Sin embargo, la suposición de que los errores tendrán la misma dispersión en todas las clases quizás sea excesiva. Empíricamente se ve que normalmente las magnitudes de los residuos dependen de las magnitudes de las densidades de los intervalos.

Ajuste de la Mínima JI cuadrado.
Obtener los parámetros θ del espacio paramétrico ω tales que hagan mínima la expresión:

$$\sum_{i=1}^{nc}\frac{1}{PY_i}(F_i - PY_i)^2$$

Este método pondera pues más las clases con probabilidades pequeñas que lo hace el de Máxima Verosimilitud, por lo que las estimaciones de mínima χ^2 pueden ajustar mejor los extremos de la distribución.

4.4 Consideraciones prácticas para la selección de un modelo

La literatura existente nos alerta sobre la posibilidad de alta correlación entre los estimadores de los parámetros obtenidos por los distintos métodos (p.e. Johnson y Kotz aseguran una alta correlación entre los parámetros α y c de la Distribución Gamma generalizada). Para estos casos hemos previsto la utilización de relaciones robustas entre los datos y los parámetros, propuestas por Crow, Siddiqui, Raghunandana, Robson y Whitlock.

Para medir la bondad del ajuste alcanzado, el principal obstáculo está según Selvin y Stuart en

1966, en que si utilizamos los datos para determinar un modelo concreto de distribución, invalidamos cualquier test de ajuste subsecuente realizado sobre los mismos datos. Así pues, necesitaremos algún indicador que resuma la bondad del ajuste de cada distribución por comparación con sus actuales frecuencias. Y para ello, la solución inmediata consiste en tomar como medida de bondad del ajuste la medida de discrepancia entre modelo teórico y distribución empírica que proporciona directamente el criterio a minimizar utilizado por cada uno de los métodos de estimación empleados.

Hemos resumido los resultados obtenidos en la siguiente Tabla, donde figura la estadística resultante para las distribuciones que sistemáticamente vienen presentando las primeras posiciones. (posiciones en cada caso, una vez ordenados de mejor a peor ajuste, y media y desviación típica de las posiciones).

Así, tras presentar el número de orden que han ocupado en cada ocasión, habiendo eliminado de la tabla las distribuciones Dagum I, II y III, al considerar la mejor estimación resultante para esta familia como aquella que se deriva de estimar el mejor modelo de tipo II ó III a partir de la solución inicial dada por el ajuste de la de tipo I, y que hemos representado simbólicamente como F. Dagum (Familia de Dagum), aparece asociado a cada una de estas distribuciones el lugar promedio ocupado así como la desviación típica de tales lugares en la ordenación, con el fin de apreciar no sólo el posicionamiento general de ésta, sino la estabilidad de la misma en la clasificación.

Examinados los resultados que para cada encuesta considerada, y para las distintas distribuciones cuyo ajuste ha sido intentado por cada uno de los cuatro métodos empleados, se observa una constancia clara de las distribuciones de Dagum, Log-Student, Log-Logística (límite de la distribución de Champernowne) y Log-Normal a estar en las primeras posiciones. De las cuatro distribuciones que mejores resultados proporciona, en el año 73 el primer puesto se lo disputan entre las distribuciones de Dagum, Log-Logística y Log-Student, seguidas lógicamente de la distribución de Champernowne con el parámetro O casi nulo (límite de la Logística), y a continuación la Log-Normal. En el año 80, los primeros puestos son para las distribuciones Log-Student, Log-Normal, Dagum, Log-Logística y Champernowne.

Del análisis de estas tablas deducimos en primer lugar, con independencia del método de ajuste empleado, que las dos distribuciones que sistemáticamente suelen acaparar las primeras posiciones son las correspondientes a la Familia de Dagum seguida a continuación de la distribución Log-Student (solamente se invierte este orden cuando el método de ajuste utilizado fue el de mínimos cuadrados a la función de densidad), y en todos ellos la estabilidad de la posición obtenida es mucho más consistente en el caso de la Familia de Dagum que la de la Log-Student, al observarse desviaciones típicas sistemáticamente menores para todos los métodos de estimación empleados.

En segundo lugar, los métodos de estimación que mayor regularidad o estabilidad proporcionan a las ordenaciones obtenidas son por este orden, el de máxima verosimilitud, seguido del de la mínima χ^2, los cuales preservan las primeras posiciones para las distribuciones que teóricamente hemos visto que presentan mejores propiedades lógico-empíricas, y cuya utilización se muestra más fundamentada desde un punto de vista económico.

Resultados según los distintos métodos de ajuste.

Método:	Distribuciones en función de su bondad de ajuste	Datos Básicos			Datos corregidos: Tasa Ocult.constante			Datos corregidos: Tasa Ocult. progresiva			Media	Desv. Típica
		73/74	80/81	90/91	73/74	80/81	90/91	73/74	80/81	90/91		
Máxima Verosimilitud	LOG-STUDENT	1	4	3	7	2	4	2	1	2	2.89	1.79
	FAM.DAGUM	2	3	2	1	4	3	1	2	4	2.44	1.07
	SINGH-MADDALA	3	1	1	2	1	2	6	3	3	2.44	1.50
	CHAMPERNOWNE	4	5	5	3	6	5	3	5	6	4.67	1.05
	LOG-LOGISTICA	5	6	6	4	5	6	5	6	5	5.33	0.67
	BOX-COX-CHAMPERNOWNE	6	2	4	5	3	1	4	4	1	3.33	1.63
	LOG-NORMAL	8	9	9	8	9	9	8	8	9	8.56	0.50
Mínima χ^2	LOG-STUDENT	1	6	3	2	1	3	2	1	1	2.22	1.55
	FAM.DAGUM	2	2	2	1	2	2	1	3	3	2.00	0.67
	SINGH-MADDALA	3	1	1	3	5	1	3	2	2	2.33	1.25
	CHAMPERNOWNE	20	5	18	13	14	9	16	13	14	13.56	4.25
	LOG-LOGISTICA	4	3	4	4	3	4	4	4	4	3.78	0.42
	BOX-COX-CHAMPERNOWNE	21	15	11	17	18	12	7	20	19	15.56	4.42
	LOG-NORMAL	6	8	7	6	8	7	6	6	7	6.78	0.79
Mínimos Cuadrados a la F. de Distribución	LOG-STUDENT	1	2	2	1	1	2	2	15	1	3.00	4.27
	FAM.DAGUM	2	1	1	2	3	1	1	1	2	1.56	0.68
	SINGH-MADDALA	13	6	5	5	2	3	7	7	5	5.89	2.96
	CHAMPERNOWNE	6	4	4	4	6	4	17	4	4	5.89	4.01
	LOG-LOGISTICA	4	3	3	3	7	5	3	3	3	3.78	1.31
	BOX-COX-CHAMPERNOWNE	5	19	26	18	4	6	15	2	20	12.78	8.15
	LOG-NORMAL	3	5	6	6	8	7	4	6	6	5.67	1.41
Mínimos Cuadrados a la Función de Densidad	LOG-STUDENT	1	2	1	2	8	1	1	1	1	2.00	2.16
	FAM.DAGUM	5	3	2	4	2	3	2	2	2	2.78	1.03
	SINGH-MADDALA	7	13	6	14	6	2	13	4	3	7.56	4.35
	CHAMPERNOWNE	8	6	5	7	4	6	6	5	6	5.89	1.10
	LOG-LOGISTICA	9	5	4	6	5	5	7	7	4	5.78	1.55
	BOX-COX-CHAMPERNOWNE	10	4	3	8	3	4	8	6	5	5.67	2.36
	LOG-NORMAL	6	9	9	5	9	7	3	8	8	7.11	1.97

Como conclusión, teniendo en cuenta que para la distribución de Dagum el parámetro a puede considerarse prácticamente nulo, podemos decir que 2 ó 3 parámetros son suficientes para ajustar adecuadamente las distribuciones de la renta per cápita en España y en estos años. Y si bien la fundamentación económica nos inclinaría a considerar como modelos estables en el tiempo los de Dagum y Log-Logística (elegir Champernowne sería en cierta forma redundante con la elección de la Log-Logística dado el valor que toma su parámetro O), la simplicidad de las leyes Log-Student y Log-Normal unida a su buen comportamiento, las hace atractivas a la hora de ser consideradas como modelos eficientes. Sin embargo, puesto que la mayor ocultación parece darse en las clases altas de ingresos, una vez corregidos los datos en este sentido, probablemente estas últimas distribuciones empeoren algo su ajuste, al exigírseles un menor grado de contacto en estas clases con el eje de abscisas, por lo que en nuestra opinión, hasta ver el comportamiento que presentan con la nueva EPF de 1990, y repetir finalmente el proceso para los datos corregidos por ocultación, no debemos descartar ninguno de ellos.

Por estos motivos, y teniendo en cuenta tanto las propiedades teóricas como empíricas de los métodos y de los modelos en cuestión, concluimos que como mejor método de estimación, tanto por sus propiedades teóricas conocidas, como por su comportamiento empírico

observado, en tanto que proporciona los mejores ajustes para los mejores modelos probabilísticos, preservando una gran estabilidad para las ordenaciones de los mismos, hemos adoptado como procedimiento de ajuste en el resto del trabajo, el método de máxima verosimilitud.

Como modelo probabilístico más adecuado a los datos observados, en base a los resultados obtenidos, consideramos por sus propiedades teóricas y empíricas a la Familia de Dagum. En cuanto a propiedades de tipo empírico, se presenta casi a la par el modelo Log-Student, lo cual hace prever un buen ajuste de tipo ad-hoc a los datos, si bien la interpretabilidad de sus parámetros y su fundamentación de tipo económico dejan mucho que desear. De todas formas, dada la bondad del ajuste que presenta, hemos decidido en una primera instancia seguir el estudio con los dos modelos citados, a fin de observar su adecuación a las demás categorías que van a ser estudiadas: regiones, categorías socioprofesionales, y clases de hábitat, pero finalmente fue la Familia de Dagum la que presentó mejores resultados de una forma más general, y por ello definitivamente seleccionada.

4.5 Resultados de la modelización

Observando los resultados de los ajustes efectuados con estas selecciones, y presentados en la tabla del Anexo 1, bajo la columna *Modelo Teórico*, podemos apreciar en una primera instancia que, si bien los resultados obtenidos son bastante comparables con los empíricos, en general, se observa a veces pequeñas discrepancias en los valores de las medias poblacionales, por lo que, puesto que el número real de perceptores es constante y conocido en cada año, el monto resultante de renta repartida entre ellos, no coincide exactamente con las cantidades de referencia que fueron deducidos de acuerdo con los datos de la Contabilidad Nacional. Por ello hemos planteado el problema de ajustar los datos a la Familia de Dagum, una vez que la hemos seleccionado para nuestro estudio global, pero imponiendo una tal condición de armonización con aquellos datos de referencia.

Para ello, puesto que el valor teórico de la media de la distribución de Dagum tiene una expresión explícita, hemos introducido en el procedimiento de optimización que nos permite obtener el ajuste de máxima verosimilitud, la restricción adicional que resulta de igualar la media teórica a la media empírica deducida, y que queremos respetar; volviendo a repetir el proceso de ajuste mediante este procedimiento, para las distintas categorías estudiadas (clases de hábitat, categorías socioprofesionales, regiones y nivel nacional), y los distintos años (73/74, 80/81 y 90/91), y cuyos resultados se presentan asimismo en la Tabla del Anexo 1, bajo la columna *Modelo Armonizado*.

Una primera mirada comparativa a esta tabla, nos hace apreciar un alto grado de aproximación entre los datos empíricos y los teóricos, si bien esta aproximación es mayor en el caso del modelo teórico armonizado. Para éste, y como ya se dijo, se impuso la condición de que el valor medio coincidiera con el observado empíricamente, lo que hace que también coincidan las cantidades correspondientes a la renta total agregada, el índice de posicionamiento con respecto de la nacional de la renta media y el porcentaje de participación en el total nacional, y que se aproximen más la práctica totalidad de los datos recogidos en la tabla, en la práctica totalidad de los casos.

Cabe sin embargo y a modo general dos observaciones. La primera se relaciona con la discrepancia existente entre las distintas rentas máximas presentadas. Recordemos que la renta

máxima obtenida empíricamente era la mayor de las rentas observadas, la cual venía afectada de un factor de elevación debido al muestreo, lo que implicaba la representación en sí de un número relativamente grande de personas. En el caso de los modelos teóricos, ésta se estima como la mínima renta para la que no cabe esperar ninguna persona con una renta mayor. Esto explica obviamente la diferencia, y además pensamos que por este motivo los modelos teóricos presentan una estimación más realista de estas rentas máximas.

La segunda se refiere a las cantidades relacionadas con los momentos de ordenes mayores o iguales a 2 (desviación típica, coeficiente de variación de Pearson y coeficiente de simetría). La idea de buscar los modelos más apropiados para representar distribuciones de rentas, entre aquéllos que poseen un número finito de momentos finitos, se refuerza al observar los valores empíricos que ponen de manifiesto estas tablas: coeficientes de variación bastante superiores al 100% de las medias y coeficientes de simetría de varias decenas de magnitud. Los modelos teóricos ajustados no poseen momentos de tercer orden, por lo que no es posible calcular los correspondientes coeficientes de simetría. Sin embargo, y a pesar de todo, algunas veces, parece que el grado de contacto de la cola superior de la distribución con el eje de abscisa debería ser algo menor que el que provee estas distribuciones ajustadas, ya que en algunas ocasiones, la desviación típica se ve subestimada, con la consiguiente subestimación asimismo del coeficiente de variación de Pearson.

Por lo demás, el grado de aproximación parece bastante satisfactorio en el resto de cantidades estimadas (incluidos los percentiles y los porcentajes de participación de cada intervalo interpercentílico en el total de la renta repartida), incluso en el caso del intervalo que mayor discrepancia suele presentar y que siempre es el último de la cola derecha (p95-Max), para el que no suele registrarse un error de participación superior al 1% del total, y sólo en algunos casos realmente anómalos, puede mostrarse superior a un 2 ó 3 por ciento.

5 ESTUDIO DE LA DESIGUALDAD.

El análisis de la desigualdad se desarrolla mediante una triple aproximación, que permita comparar y ordenar las distintas distribuciones de renta generadas por las clasificaciones que se desea utilizar (global, Comunidades Autónomas, Categorías Socio-Profesionales y Tipos de Hábitat). Esta triple aproximación viene determinada por la utilización de las curvas de Lorenz, las curvas generalizadas (Shorrocks) y los indicadores de desigualdad y de nivel de vida-renta[1]. Conviene enfatizar que lo que se pretende no es medir el bienestar social, sino tan solo uno de sus componentes, que es la desigualdad de rentas y, de ningún modo, se pretende identificar ambos conceptos.

5.1 Curvas de Lorenz.

Se define como la relación existente entre la proporción acumulada de perceptores de renta y la proporción correspondiente de rentas recibidas, cuando dichos perceptores se ordenan en sentido creciente en relación con su renta. Así pues, permite estudiar la

[1] Entendemos por *nivel de vida-renta* el componente del nivel de vida de los individuos desde el punto de vista exclusivo de la renta. Esto excluye, por tanto, otras componentes de bienestar relacionadas con la teoría de la utilidad o las funciones de bienestar social.

desigualdad del reparto de la distribución de la renta, mediante la comparación con la situación de reparto igualitario en la que todos los individuos obtendrían la misma renta (la renta media), sin que esto signifique postura alguna acerca de la justicia social de tal tipo de reparto, ya que se utiliza sólo como patrón de referencia. Por lo tanto, suponiendo que el vector de rentas de la población es (x_1, x_2,..., x_k), cuyas componentes están ordenadas de menor a mayor, y que (f_1, f_2,..., f_k) es el correspondiente vector de frecuencias relativas, entonces:

$$p_i = \sum_{j=1}^{i} f_j$$

$$q_i = L(p_i) = \frac{1}{\mu} \sum_{j=1}^{i} x_j . f_j$$

y la curva de Lorenz queda definida por los puntos $\{(p_i,q_i), i=0,1,..., k\}$, donde μ es la renta media y $p_0 = q_0 = 0$. Su representación gráfica queda inscrita en el cuadrado unidad, siendo su diagonal principal la correspondiente a un reparto igualitario de rentas. Por tanto, a efectos de comparación, una distribución de rentas será más desigual que otra si la curva de Lorenz asociada con aquella encierra completamente a la de ésta. Esta operación es la que define el *criterio de dominación de Lorenz*:

$$x \leq_L y \Leftrightarrow L_x(p) \leq L_y(p), \forall p \in [0,1]$$

No obstante, esta relación no permite una ordenación completa, ya que frecuentemente se presentan intersecciones entre las curvas de Lorenz que se comparan. Esto hace que la utilización de este criterio conduzca a cuasi-ordenaciones en las que abundan las parejas de distribuciones no comparables pero que, no obstante, proporcionan una información muy útil.

5.2 Curvas de Lorenz Generalizadas o de Shorrocks.

Las limitaciones de la utilización del criterio de Lorenz impiden relacionar ordinalmente una gran cantidad de casos. Como intento de solución, se ha propuesto la utilización de la curva de Lorenz generalizada, introduciendo un cambio de escala a través de la renta media:

$$LG(p_i) = \mu . L(p_i)$$

que conduce al *criterio de dominación generalizado*, obtenido de modo similar:

$$x \leq_{LG} y \Leftrightarrow LG_x(p) \leq LG_y(p), \forall p \in [0,1].$$

No obstante, proporciona igualmente una cuasi-ordenación, aunque ahora el número de comparaciones no resueltas, por la aparición de intersecciones, es notablemente inferior. Ahora bien, la utilización de este criterio admite implícitamente que la renta media de una población incide sobre la comparación, ya que supone admitir que la sociedad prefiere repartos con mayor renta media para niveles similares de desigualdad, por lo que su interpretación debe efectuarse en términos de nivel de vida-renta y no en términos de desigualdad.

Puesto que estas curvas son de tipo absoluto y vienen definidas en la misma unidad que la renta, resulta obvio que la comparación inter-temporal deberá realizarse con las unidades monetarias convenientemente deflacionadas.

5.3 Medidas de Desigualdad.

Los criterios de dominación permiten un estudio global comparativo de la desigualdad y el nivel de vida-renta para las distribuciones de renta. Sin embargo, para cuantificar estos conceptos, es necesaria la utilización de indicadores de desigualdad y de nivel de vida-renta. Ahora bien, estas medidas permiten la ordenación total de diferentes dstribuciones de renta, mediante el uso de indicadores sintéticos, por lo que diferentes indicadores generarán ordenaciones diferentes, ya que sus formulaciones llevarán implícitamente incorporados distintos esquemas de ponderación sobre cada uno de los tramos de renta. Por ello, la selección de estas medidas deberá hacerse analizando la validez de los sistemas de ponderación, lo que conduce a la implantación de una serie de axiomas que garanticen que el comportamiento de las medidas resulte coherente.[2]

Para efectuar la selección de las medidas de desigualdad que se utilizarán, un buen punto de partida consiste en agrupar por familias un amplio conjunto de indicadores utilizados en otros estudios previos. Entre estas familias, aparte de las medidas basadas en estadísticos descriptivos básicos, pueden citarse las de Atkinson, de Theil, las relacionadas con la diferencia media (entre ellas, destaca el Indice de Gini), las medidas de divergencia basadas en distancias entre distribuciones y las medidas de divergencia intercuantílica.[3] De esta manera, se obtuvo un conjunto integrado por 38 indicadores, de los que la selección final propuesta consta de 8 medidas mediante la aplicación de dos criterios.

En primer lugar, se estudiaron las relaciones existentes entre los indicadores mediante transformaciones monótonas, que preserven la ordenación final proporcionada; así, por ejemplo, el indicador de Theil de orden 0 está relacionado con el de Atkinson de orden 1, mediante la transformación $f(x) = 1-\exp(-x)$, lo que ha permitido elegir representantes de distintas *clases* cuyos integrantes están relacionados mediante este tipo de transformaciones.

En segundo lugar, se ha analizado la serie de axiomas que verifican los indicadores resultantes. Estos axiomas se han seleccionado de manera que las ordenaciones resulten compatibles con la proporcionada por el criterio de dominación de Lorenz, añadiendo algunos más de uso frecuente en los estudios de desigualdad. En concreto, dichos axiomas son los siguientes:

A.1.- *Simetría* (SIM).- La desigualdad es invariante frente un cambio de orden o permutación en las rentas de la distribución.

A.2.- *Invarianza frente a cambios de escala* (INVAR).- El indicador se mantiene invariante frente a variaciones proporcionales en todas las rentas; es decir, $I(\lambda x)=I(x)$, $\forall\lambda>0$.

---

[2] Se adopta aquí el enfoque *positivo*, que consiste en la consideración meramente objetiva de las medidas, frente al enfoque *normativo*, relacionado con valoraciones de tipo ético en relación con el bienestar social aportado por las rentas.

[3] La definición concreta de estas familias puede verse, por ejemplo, en Bartels(1977) y en Callealta, Casas y Núñez (1994).

A.3.- *Principio de Población de Dalton* (POBL).- El indicador debe mantenerse invariante frente a réplicas (repetición un mismo número de veces de cada renta) de la población.

A.4.- *Principio de Transferencias de Pigou-Dalton* (TRANSF).- Si se produce una transferencia de renta de un individuo hacia otro con menor renta (transferencia progresiva), de manera que deje inalterada la ordenación inicial, la desigualdad debe decrecer.

A.5.- *Normalización* (NORM).- El indicador de desigualdad debe tomar valores en el intervalo [0,1], correspondiendo el 0 al caso de equidistribución y el 1 al de máxima desigualdad.

A.6.- *Descomposición Aditiva en Subgrupos* (DESCOMP).[4]- Si la población se divide en subgrupos, el indicador puede separarse de manera aditiva en términos de la desigualdad entre grupos y la encontrada dentro los mismos. Es decir:

$$I(x) = I_e(x_{(1)},...,x_{(r)}) + \sum_{h=1}^{r} \alpha_h I(x_{(h)})$$

Finalmente, se han eliminado también aquellos indicadores que presentan esquemas de ponderación muy irregulares e irreales, quedando reducida la selección a los ocho casos aludidos previamente. Estos indicadores, ya expresados en su versión estocástica para facilitar su uso a través del modelo probabilístico, son los siguientes, siendo μ la renta media:

a) Típicamente de dispersión:

- Coeficiente de Variación Cuadrado Normalizado:[5] $CV2.NORM = CV^2/(1+CV^2)$
- Varianza Normalizada de Logaritmos: $VL.NORM = VL/(1+VL)$

b) Explícitamente de Desigualdad:

- Indice de Gini: $GINI = E[|X-Y|]/(2\mu)$; es decir, la diferencia media estandarizada y normalizada.
- Indice de Theil Normalizado de orden 1:
  $TH1.NORM = 1- \exp(-T1)$, siendo $T1 = E[X.\log(X/\mu)]/\mu$
- Indice de Atkinson de orden 0.5, con escasa aversión a la desigualdad[6]:
  $ATKIN0.5 = 1 - [E(X^{1/2})]^2/\mu$
- Indice de Atkinson de orden 1, con una aversión a la desigualdad de tipo medio:
  $ATKIN1 = 1 - \exp[E(\log(X/\mu)]$, relacionado con la media geométrica.
- Indice de Atkinson de orden 2, con una alta aversión a la desigualdad:
  $ATKIN2 = 1 - (E[1/X])^{-1}/\mu$, relacionado con la media armónica.
- Indice de Pietra ó Porcentaje de Máxima Igualación:
  $PIETRA = E[|X-\mu|]/(2\mu)$

---

[4] Se considera la descomposición aditiva, en el sentido expuesto en Bourguignon (1979) ó Shorrocks (1980). Existen, sin embargo, esquemas de descomposición más complicados.

[5] Las versiones normalizadas se obtienen mediante la utilización de transformaciones estrictamente crecientes que apliquen el rango del indicador en [0,1].

[6] La familia de índices de Atkinson puede obtenerse a partir de la función general de promedios, utilizando distintos valores que se interpretan como un parámetro de aversión a la desigualdad.

El cuadro de análisis de los axiomas que verifican los indicadores seleccionados, figura a continuación:

TABLA 5.1.- Verificación de los axiomas propuestos, según los distintos indicadores.

	TRANSF.	INVAR.	POBL.	SIM.	NORM.	DESCOMP.
CV2.NORM.	SI	SI	SI	SI	SI	SI
VL.NORM.	NO	SI	SI	SI	NO	SI
GINI	SI	SI	SI	SI	SI	NO
TH1.NORM.	SI	SI	SI	SI	SI	SI
ATKIN. 0.5	SI	SI	SI	SI	SI	NO
ATKIN. 1	SI	SI	SI	SI	SI	NO
ATKIN. 2	SI	SI	SI	SI	SI	NO
PIETRA	NO	SI	SI	SI	SI	NO

A partir de estos indicadores, pueden obtenerse los correspondientes indicadores de nivel de vida-renta, sin más que efectuar la transformación:

$$\text{INV-R} = \mu.(1\text{-DES}),$$

donde DES representa el indicador de desigualdad elegido, y μ es la renta media, expresada en unidades monetarias constantes, para facilitar las comparaciones inter-temporales, en este caso del año 1986. De esta forma, se observa el comportamiento esperado en este tipo de indicadores; es decir, aumentan en relación directa con la renta media e inversa con el nivel de desigualdad.

El uso de una batería de indicadores como la propuesta permite definir otra cuasi-ordenación, que elimine, en parte, la subjetividad implícita en la elección de un único indicador, ya que cada uno pondrá el énfasis en tramos concretos de la distribución. Por ello, si se admite la tesis de A. Sen[7], para el que *la desigualdad es, en esencia, una cuasi-ordenación,* se podría utilizar una relación del tipo que el mismo autor propone, en el sentido de definir la relación *x es, al menos, tan desigual como y* (xRy), entre dos distribuciones de renta, del siguiente modo:

$$xRy \Leftrightarrow I_i(x) \geq I_i(y), \quad \forall i$$

que determinará, de nuevo, una cuasi-ordenación similar a la obtenida mediante el criterio de dominación de Lorenz, ya que los indicadores seleccionados son compatibles con esta relación, excepto quizás VL.NORM, que no cumple el Principio de Transferencias en niveles altos de renta.[8].

Obviamente, este análisis no elimina la posibilidad de cuantificar la desigualdad mediante una medida concreta, siempre que se admitan las implicaciones normativas derivadas de su uso, determinadas a partir de los axiomas que verifica, analizando el esquema de ponderaciones que utiliza y la distribución de rentas de referencia sobre la que trabaja, entre otros aspectos. En este sentido, pueden considerarse básicos los indicadores de Gini y Pietra, por su clara interpretación en términos de desigualdad.

---

[7] Ver, por ejemplo, Sen (1973).

[8] Esta cuasi-ordenación puede refinarse más aún, eliminando del criterio aquellas medidas que presentan un comportamiento más irregular con respecto al resto, para obtener resultados más concluyentes, renunciando a los esquemas de ponderación de los indicadores suprimidos y aceptando sólo los del resto.

5.4 - El caso de España en el Periodo 1973-1991.

Como ya se ha expuesto, los datos que configuran las distribuciones de renta analizadas corresponden a rentas *per cápita*, convenientemente corregidas del efecto de ocultación, y provenientes de las Encuestas Básicas de Presupuestos Familiares de 1973/74, 1980/81 y 1990/91. Así pues, trabajamos con muestras convenientemente seleccionadas; no obstante, puede ocurrir que se produzcan comportamientos anómalos en los percentiles extremos, por lo que muchos estudios suelen trabajar directamente con los deciles de la distribución, lo que se agudiza si se ha efectuado alguna clasificación y se haya provocado una fuerte desagregación. Por ello, pese a construir las curvas de Lorenz y sus generalizadas en base a los percentiles, se impone una cierta cautela al tener en cuenta los cruces de curvas que se produzcan en los extremos. Esto ha llevado a la determinación de construir las relaciones de dominación a partir de las curvas de Lorenz *truncadas*, e igual tratamiento para las curvas generalizadas. Por lo tanto, el criterio de Lorenz se expresaría ahora en los siguientes términos:

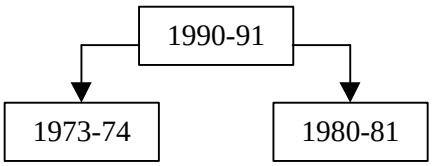
$$x \leq_L y \iff L_x(p) \leq L_y(p), \forall p \in [0.05, 0.95]$$

Con esta modificación, el número de comparaciones que dan lugar a relaciones de dominación está más acorde con los resultados obtenidos en otros estudios de este tipo. El mismo problema de estabilidad presentan las medidas CV2.NORM y VL.NORM, al hacer recaer fuertes ponderaciones sobre las rentas más extremas. Este hecho se pone de manifiesto cuando se comprueban los resultados muestrales con los obtenidos a partir del modelo probabilístico ajustado, ya que, en este caso, la estabilidad que presentan estas dos medidas es muy superior. El resto de los indicadores sí presentan un comportamiento mucho más estable, al no ponderar las rentas extremas con tanta incidencia.

En relación con los resultados muestrales obtenidos, se presentan los resultados derivados del estudio longitudinal realizado para el caso de la distribución de la renta española, durante el periodo 1973-1991, así como un caso como muestra de los diversos análisis transversales desarrollados; en concreto, el análisis comparativo de la desigualdad para las diversas Comunidades Autónomas, en el año 1990-91.

5.4.1 Análisis de la desigualdad de España para el periodo 1973-1991.

La comparación de las curvas de Lorenz correspondientes puede sintetizarse mediante el siguiente esquema, donde el sentido de las flechas indica aumento de desigualdad:



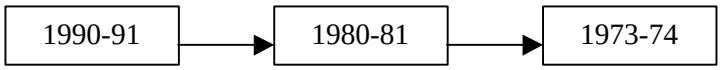
Así pues, observamos como la desigualdad ha disminuido de manera global en todo el periodo, aunque se registra un punto de corte entre las curvas de 1973-74 y 1990-91, que se sitúa por debajo del percentil 3, con lo que puede aplicarse la argumentación ya mencionada. En cuanto a la evolución entre 1973-74 y 1980-81, se observa una disminución general de la desigualdad, excepto en las rentas extremas; en concreto, por debajo del primer decil y por encima del percentil 98. Los resultados de la batería de índices seleccionados se presentan a

continuación, donde las conclusiones son similares de manera global, aunque las diferentes ponderaciones provocan algún cambio en casos aislados.

TABLA 5.2.- Resultados de los indicadores de desigualdad en España, en el periodo 1973-91.

AÑOS	CV2.NORM	PIETRA	GINI	TH1.NORM	VL.NORM	ATKIN2	ATKIN0.5	ATKIN1
1973/74	0,5900	0,2717	0,3866	0,2590	0,3081	0,3732	0,1256	0,2214
1980/81	0,7701	0,2627	0,3757	0,2716	0,3031	0,3797	0,1234	0,2153
1990/91	0,6472	0,2439	0,3496	0,2233	0,2787	0,3393	0,1052	0,1872

Si se utilizan las curvas generalizadas de Lorenz, para analizar la evolución del nivel de vida-renta, los resultados, en términos de dominación, se resumen en el siguiente esquema:



Se deduce, por tanto, que se ha producido un incremento en términos de nivel de vida-renta de manera continuada durante el periodo analizado.

5.4.2 *Análisis de la desigualdad en España, según Comunidades Autónomas, en 1990-91.*

Utilizando el criterio de dominación de Lorenz para curvas truncadas, puede establecerse el esquema de dominación que se presenta en el Gráfico 5.1, donde el sentido de las flechas indica mayor desigualdad, mientras que dos Comunidades Autónomas no conectadas indican que no puede establecerse una relación de dominación entre ellas, por haberse producido cruces entre sus respectivas curvas de Lorenz. El porcentaje de comparaciones que han generado ordenaciones asciende al 55,56%, mientras que sólo habría sido del 31,37%, si se hubiese utilizado el criterio de Lorenz en términos rigurosos.

La dificultad para trazar el esquema completo obliga a su partición en dos subesquemas, de manera que el segundo presenta las relaciones de dominación que afectan de modo directo a la Comunidad Autónoma de Castilla-León, que no pudo incluirse en el primero. Por otra parte, la Tabla 5.3 muestra las ordenaciones proporcionadas por las distintas medidas de desigualdad, donde el 1 indica la mayor desigualdad y el 18 el mínimo. Puede comprobarse como, en términos generales, se respeta la ordenación que proporciona el esquema con escasas discrepancias, cuyas causas ya se han comentado, no obstante.

Por último, se presenta el esquema de dominación generado por las curvas de Lorenz generalizadas y truncadas, lo que ahora permite obtener un 77,12% de comparaciones válidas. La diferencia con el análisis anterior estriba en que ahora la dominación no se establece en términos de desigualdad sino de nivel de vida-renta.

GRÁFICO 5.1.- Esquema de dominación de Lorenz, para el año 1990-91.

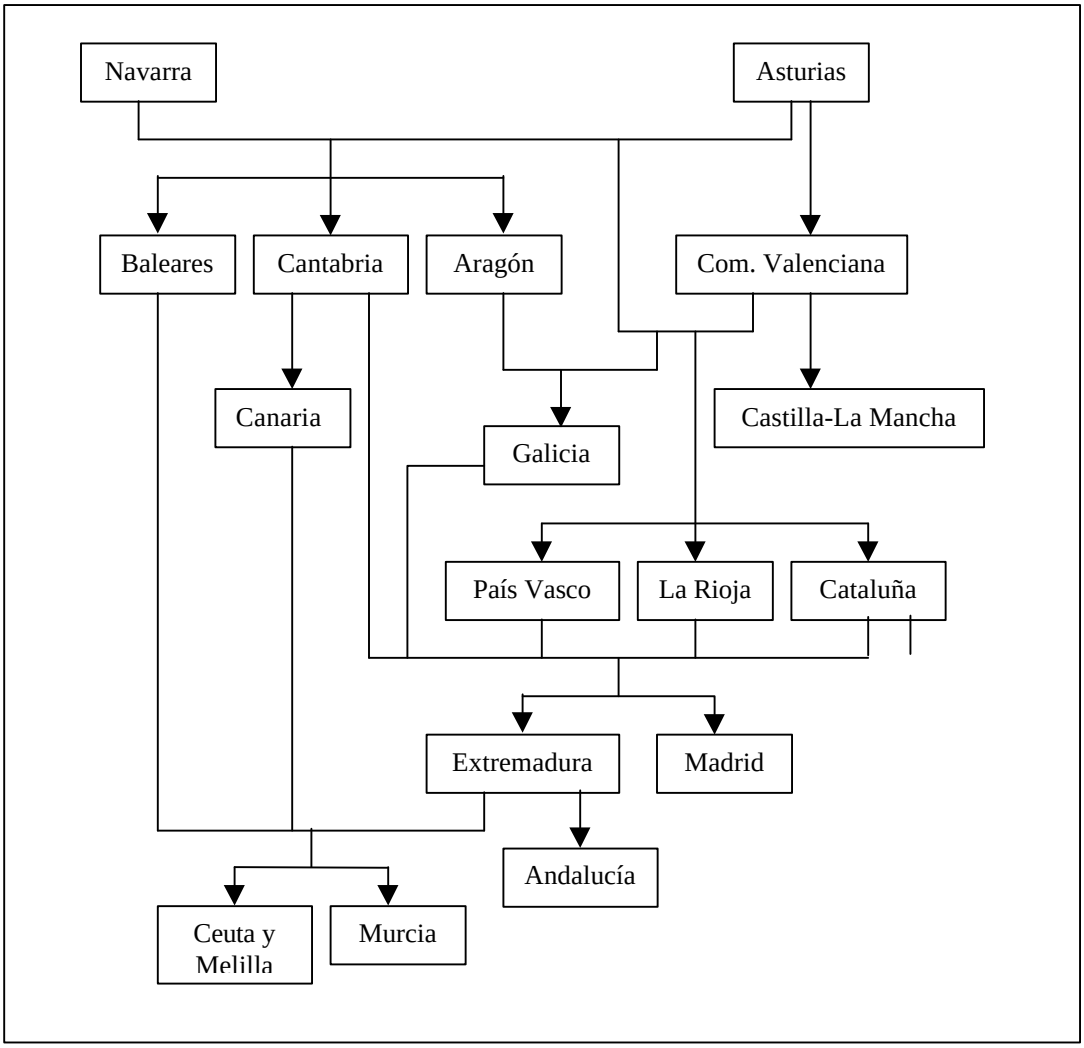


GRÁFICO 5.1.- Esquema de dominación de Lorenz, para el año 1990-91.(Cont.)

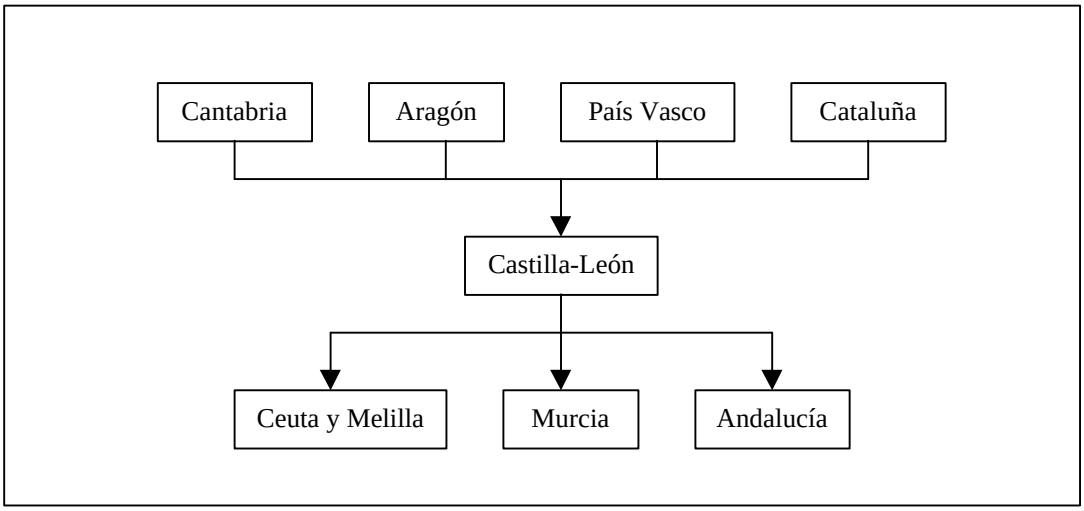


TABLA 5.3.- Ordenaciones proporcionadas por las medidas de desigualdad. Año 1990-91.

	CV2.NOR.	PIETRA	GINI	TH1.NOR.	VL.NOR.	ATKIN2	ATKIN0.5	ATKIN1
Andalucía	3	4	4	4	4	4	5	3
Aragón	8	16	16	14	14	13	16	16
Asturias	17	18	18	18	18	18	18	18
Baleares	16	10	12	15	5	5	13	11
Canarias	14	6	6	8	3	3	6	6
Cantabria	15	14	14	16	8	7	14	12
Castilla-León	11	7	8	7	7	9	8	8
Castilla-La Mancha	1	5	5	1	16	10	1	5
Cataluña	13	9	9	12	10	11	11	9
Com. Valenciana	12	15	15	13	15	14	15	15
Extremadura	10	8	7	6	6	8	7	7
Galicia	5	12	11	9	12	15	9	13
Madrid	2	3	3	2	9	6	3	4
Murcia	4	2	2	3	2	2	4	2
Navarra	18	17	17	17	17	17	17	17
País Vasco	7	11	10	10	11	12	10	10
La Rioja	9	13	13	11	13	16	12	14
Ceuta y Melilla	6	1	1	5	1	1	2	1

GRÁFICO 5.2.- Esquema de dominación generalizado de Lorenz. Año 1990-91