



Working Paper

Àrea d'Economia

Àlex Costa
Vittorio Galletto
Jaume Garcia
Josep Lluís Raymond
Daniel Sánchez-Serra

**Pobreza y desigualdad en las
Comunidades Autónomas.
Análisis de los resultados de la Encuesta
de Condiciones de Vida y del Atlas de
Distribución de Renta de los Hogares:
una propuesta de distribución
combinada**

Institut Metròpoli Working Paper in Economics, nº 24.01



Institut Metròpoli Working Paper in Economics, n° 24.01 October 2024

Pobreza y desigualdad en las Comunidades Autónomas.
Análisis de los resultados de la Encuesta de Condiciones de Vida y del Atlas de Distribución de Renta de los Hogares: una propuesta de distribución combinada

Autores: Àlex Costa, Vittorio Galletto, Jaume Garcia, Josep Luís Raymond, Daniel Sánchez-Serra

© Àlex Costa, Vittorio Galletto, Jaume Garcia, Josep Luís Raymond, Daniel Sánchez-Serra, 2024

Publicado por:
Institut Metròpoli

Parc de Recerca, Mòdul A
Universitat Autònoma de Barcelona
08193 Cerdanyola del Vallès, Barcelona, Spain.

<https://www.institutmetropoli.cat/>

Pobreza y desigualdad en las Comunidades Autónomas. Análisis de los resultados de la Encuesta de Condiciones de Vida y del Atlas de Distribución de Renta de los Hogares: una propuesta de distribución combinada¹

Àlex Costa

Departament d'Anàlisi
Oficina Municipal de Dades
Ajuntament de Barcelona
alexcostasaenz@gmail.com

Vittorio Galletto

Institut Metròpoli
Departament d'Economia Aplicada
Universitat Autònoma de Barcelona
vittorio.galletto@uab.cat

Jaume Garcia

Departament d'Economia i Empresa
Universitat Pompeu Fabra
jaume.garcia@upf.edu

Josep Lluís Raymond

Departament d'Economia i d'Història
Econòmica
Universitat Autònoma de Barcelona
josep.raymond@gmail.com

Daniel Sánchez-Serra

Directorate for Science, Technology and Innovation
Organisation for Economic Co-operation and
Development (OECD)
daniel.sanchezserra@oecd.org

Resumen: El punto de partida de este proyecto fue la disparidad de resultados entre dos fuentes del INE, la Encuesta de Condiciones de Vida (ECV) y el Atlas de Distribución de la Renta de los Hogares, para dos indicadores muy relevantes, como son el Índice de Gini, para medir la desigualdad, y Riesgo de pobreza monetaria para medir la pobreza. Se analizan estas diferencias comparando las distribuciones de ambas fuentes para la renta de los hogares por unidad de consumo. Se obtiene una distribución combinada de ambas fuentes como la función envolvente y, tomándola como referencia, se llega a la conclusión de que la ECV puede estar sobrevalorando parcialmente la pobreza e infravalorando ligeramente la desigualdad, con carácter bastante general. Por su parte, el Atlas sobrevalora parcialmente la pobreza y también sobrevalora, con carácter general, la desigualdad.

Palabras clave: *Pobreza, Desigualdad, Índice de Gini, Encuesta de Condiciones de Vida, Atlas de Distribución de la Renta*

JEL: I32, D31, D63

¹ Este documento no refleja las opiniones oficiales de las instituciones a las que pertenecen sus autores. Las metodologías y los resultados presentados son responsabilidad exclusiva de los autores. Asimismo, los autores expresan su agradecimiento al INE. Sin su apoyo este trabajo no hubiera sido posible. Los datos utilizados (siempre agregados) han estado sujetos a estrictas condiciones de confidencialidad, no divulgación y uso exclusivo para este trabajo.

1. INTRODUCCIÓN

El punto de partida de este proyecto fue la disparidad de resultados entre dos fuentes del INE, la *Encuesta de Condiciones de Vida* (ECV)² y el *Atlas de Distribución de la Renta de los Hogares* para las Comunidades Autónomas (CCAA), en dos indicadores relevantes: el Índice de Gini, que mide la desigualdad, y la pobreza en términos del Riesgo de pobreza monetaria (RdP).

Desde hace muchos años, en la estadística oficial, se han dado casos de solapamiento de fuentes sobre conceptos similares. Ejemplos conocidos de estas situaciones en la estadística española han sido los datos de paro y de ocupación, que pueden estimarse con la EPA, pero también con los registros de paro o con la afiliación a la Seguridad Social, o también los datos de población, que pueden derivarse de las operaciones censales o con registros de población, en particular con el padrón municipal.

Esta diversidad de fuentes sobre ámbitos temáticos similares ha tenido una valoración positiva, en la medida en que, por una parte, se reconoce la prioridad de los resultados de las encuestas que se ajustan mejor a los marcos conceptuales de los reglamentos y manuales internacionales, pero, por otra, se valora la capacidad de los registros para ofrecer resultados que pueden mejorar las encuestas por su puntualidad, periodicidad o detalle territorial.

Un ejemplo destacado de nuevas situaciones de solapamiento de fuentes es el caso de la Encuesta de Condiciones de Vida (ECV) y el Atlas de Distribución de la Renta de los Hogares (en adelante, el Atlas) basada en datos fiscales. Aunque los resultados presentan una buena correlación, lógicamente no coinciden.

² Para más información sobre la ECV véase: [https://www.ine.es/uc/6vpGknBui1](https://www.ine.es/uc/6vpGknBui1;); sobre el Atlas: <https://www.ine.es/uc/qxFrSJI2i1>.

A nivel internacional existen ejemplos de conciliación entre fuentes fiscales y encuestas de renta y condiciones de vida, como por ejemplo en Bélgica (De Schrijver (2020), “Fiscal data in the Statistics on Income and Living Conditions (SILC) survey: a path for the future?”, STATBL Analyse n. 08). Desarrollar una metodología que pueda facilitar esta conciliación de resultados de la ECV y del Atlas es un objetivo de este proyecto.

Un primer resultado del análisis comparativo entre las dos fuentes mostró de forma bastante clara que, aunque la fuente oficial de referencia es la ECV, existe una posible mejora de los resultados de la ECV a partir de los datos del Atlas. En una primera etapa se planteó si esa posible mejora se limitaba a los resultados de las CCAA, dejando inamovibles los del conjunto nacional (más fiables por tamaño de muestra). Pero incluso esta opción entró en crisis al constatar que también para el conjunto de España era verosímil una mejora de los resultados de la ECV utilizando los datos fiscales del Atlas.

La mejora a la que hacemos referencia es la siguiente: la cola derecha de la distribución del Atlas, donde se encuentran las rentas altas, ofrece una mayor probabilidad de que estas rentas se detecten en el Atlas que en la ECV. Este patrón es bastante sistemático, y se aprecia tanto en la distribución nacional como en las distribuciones de las CCAA.

La valoración de esta discrepancia a favor de los datos del Atlas se hizo a partir de un razonamiento económico: no existe ningún incentivo que lleve a la población con rentas altas a inflar artificialmente sus ingresos ante la administración tributaria. Si esto es así, se abre la opción a pensar que pueda existir una infracobertura para estas rentas altas en la ECV.

Por otra parte, la cola izquierda, de rentas bajas, tiene un patrón menos claro. Las diferencias entre ECV y Atlas son menores y tienen distinto signo según la comunidad autónoma. En la mayor parte de las CCAA capta más situaciones de pobreza la ECV. Una explicación puede ser la declaración de rentas que quedan fuera del registro fiscal

(economía sumergida). Pero en otras CCAA el signo es contrario y se captan más rentas bajas en el Atlas (esta es la situación de Catalunya y Madrid).

Estas valoraciones llevaron a una nueva distribución con el objetivo de utilizar lo mejor de ambas fuentes: la función envolvente. Con la función envolvente como referencia se hace una valoración de los resultados de la ECV y del Atlas en cada uno de los dos indicadores en estudio.

2. POBREZA Y DESIGUALDAD EN LAS CCAA SEGÚN ECV Y ATLAS

El punto de partida del estudio es la comparativa del indicador de pobreza en las CCAA, según la ECV y el Atlas para el año 2020 (información disponible cuando se inició el estudio). No puede sorprender que los resultados sean distintos, lo que interesa es saber si se trata de resultados significativamente diferentes. En la Tabla 1 se presentan los resultados en pobreza y desigualdad (año 2020) para las CCAA de régimen común, ciudades autónomas y su promedio³.

A partir de estos resultados nos planteamos una primera pregunta: ¿Son significativas las diferencias en los indicadores de pobreza y desigualdad para las CCAA de ECV y Atlas?

En relación con la pobreza se constata que de las 15 CCAA hay 9 en las que hay mayor pobreza según la ECV y 6 con el signo contrario. No parece un diferencial muy sistemático. Ahora bien, si atendemos a las diferencias más elevadas encontramos que las CCAA con una diferencia más significativa de pobreza captada por la ECV son: Extremadura, Andalucía, Asturias y Castilla – La Mancha. Mientras que las que tienen unos valores más elevados de pobreza según Atlas son La Rioja, Catalunya, Cantabria, Baleares y Madrid. Esta relación apuntaría a que en las CCAA con rentas más bajas capta

³ Si bien se presentan los datos de Ceuta y Melilla, debido al reducido tamaño muestral, éstas no se consideran en el análisis.

mayor pobreza la encuesta y en las de renta más alta capta más pobreza el registro fiscal (Atlas).

Tabla 1: Pobreza y desigualdad. CCAA 2020.

	POBREZA ECV	POBREZA ATLAS	DIF ECV - ATLAS	GINI ECV	GINI ECV int (*)	GINI ATLAS	DIF ECV- ECVint	DIF ECV - ATLAS	DIF ECV int - ATLAS
Andalucía	32,3	28,5	3,8	34,3	31,5	33,0	2,8	1,3	-1,5
Aragón	15,8	15,9	-0,1	28,7	26,1	29,7	2,6	-1,0	-3,6
Asturias, Principado de	20,4	16,8	3,6	31,7	28,5	30,3	3,2	1,4	-1,8
Balears, Illes	17,6	19,4	-1,8	28,5	25,4	32,8	3,1	-4,3	-7,4
Canarias	28,4	26,6	1,8	35,0	31,9	33,8	3,1	1,2	-1,9
Cantabria	15,4	17,2	-1,8	29,8	27,2	30,1	2,6	-0,3	-2,9
Castilla y León	17,9	17,6	0,3	29,5	27,1	30,2	2,4	-0,7	-3,1
Castilla - La Mancha	27,4	24,4	3,0	31,7	28,8	31,2	2,9	0,5	-2,4
Catalunya	14,8	16,6	-1,8	30,9	28,1	32,3	2,8	-1,4	-4,2
Comunitat Valenciana	25,1	23,6	1,5	32,2	29,4	32,6	2,8	-0,4	-3,2
Extremadura	32,3	28,2	4,1	30,5	27,1	31,1	3,4	-0,6	-4,0
Galicia	20,2	18,1	2,1	29,2	26,0	29,7	3,2	-0,5	-3,7
Madrid, Comunidad de	15,2	16,8	-1,6	33,4	30,4	35,9	3,0	-2,5	-5,5
Murcia, Región de	27,7	25,2	2,5	30,1	27,3	31,7	2,8	-1,6	-4,4
Rioja, La	16,1	18,6	-2,5	29,3	26,6	30,7	2,7	-1,4	-4,1
Ceuta	32,3	33,3	-1,0	39,7	37,0	42,3	2,7	-2,6	-5,3
Melilla	30,1	35,3	-5,2	41,2	38,2	42,4	3,0	-1,2	-4,2
PROMEDIO	22,9	22,5	0,4	32,1	29,2	32,9	2,9	-0,8	-3,7

Fuente: INE

(*) Elaboración propia

En relación con la desigualdad medida a través de Gini debe hacerse una observación previa que se considera muy importante. Es conocido que el índice de Gini capta más desigualdad con datos individuales que con datos agrupados, ya que la desigualdad de los datos individuales son la suma de la desigualdad entre grupos y dentro de los grupos. Ahora bien, los resultados del Atlas que publica el INE se derivan de datos agrupados que les facilita la Agencia Tributaria (en casi 200 tramos de renta). Esto hace que la comparación directa de Gini en ECV (datos individuales) y Atlas (datos agregados) no sea correcta. Lo que se ha hecho en este trabajo y se puede ver en la Tabla 1 es calcular el Gini de la ECV con los mismos intervalos de renta que el Atlas (Gini ECV int), de forma que la comparación sea homogénea.

En las últimas tres columnas de la Tabla 1 se aprecian los resultados. La diferencia entre Gini ECV (datos individuales) y Gini de ECVint es muy clara y responde a lo que se acaba de decir: para una misma fuente de información, el Gini de los datos individuales será siempre superior al de los datos agrupados (entre 2,4 y 3,4 puntos porcentuales).

Cuando se hace la comparativa homogénea, en la última columna, no hay duda de la valoración: capta más desigualdad el Atlas que la ECV, cuando en ambas se calcula el Gini con la misma agrupación. Esta mayor captación de la desigualdad se presenta sin ninguna excepción. A efectos de información complementaria se muestra en la penúltima columna la comparativa, no homogénea, entre ECV datos individuales y Atlas. El resultado es que, incluso en esta comparativa heterogénea, en conjunto son mayores los valores de Gini de Atlas que los de la ECV. Por tanto, se puede concluir que las diferencias en desigualdad entre ECV y Atlas son sistemáticas, ofreciendo mayores valores de desigualdad el Atlas.

Una vez hecha esta lectura de los resultados se plantea entender estos diferenciales a la vista de las distribuciones de la renta por unidad de consumo en cada una de estas dos fuentes.

3.- ANÁLISIS DE LA DISTRIBUCIÓN DE LA RENTA EN LAS DOS FUENTES

La información que la ECV muestra la volatilidad derivada de los errores de muestreo. Ello explica los numerosos picos observados en la función de densidad de los datos de la ECV (decv en Gráfico 1), algo que no ocurre en el caso de los datos del Atlas de distribución de la renta (datlas en Gráfico 1)⁴.

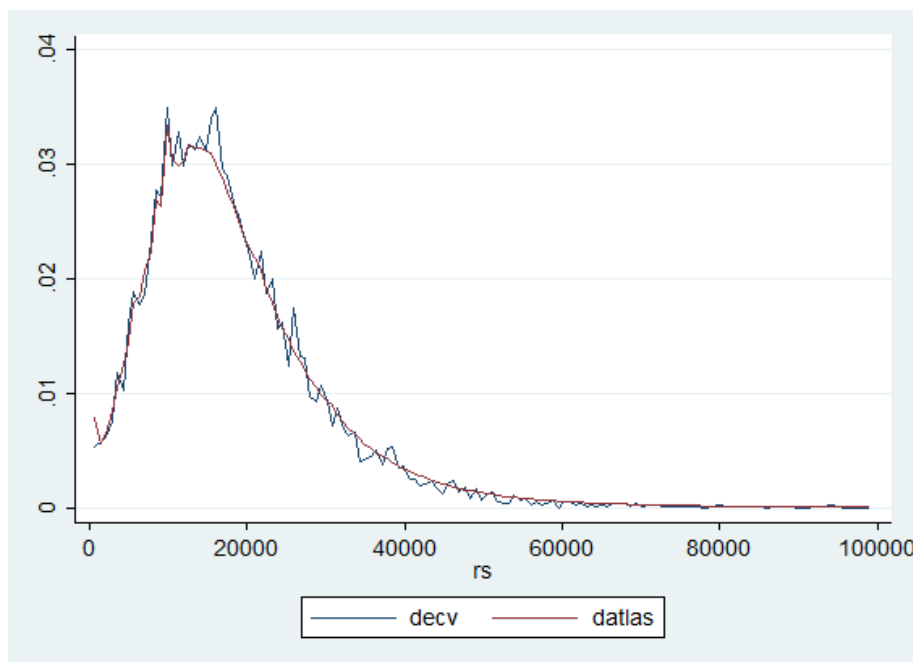
Dado que los intervalos de renta contemplados son muy reducidos, el número de observaciones que la ECV ofrece para cada intervalo suele ser muy reducido. Este error

⁴ La siguiente sección explica detalladamente la metodología seguida para la obtención de la función de densidad de probabilidades a partir del conocimiento de la función de distribución.

de muestreo es más relevante en el caso de las Comunidades Autónomas, sobre todo las pequeñas.

Para solventar este problema se ha propuesto recurrir al denominado filtro de Hodrick-Prescott (Hodrick y Prescott (1997), “Postwar US Business Cycles: An Empirical Investigation”, Journal of Money, Credit and Banking”, 29(1) 1997).

Gráfico 1. Funciones de densidad de probabilidades observadas de la ECV (decv) y del Atlas (datlas) por nivel de renta (rs)

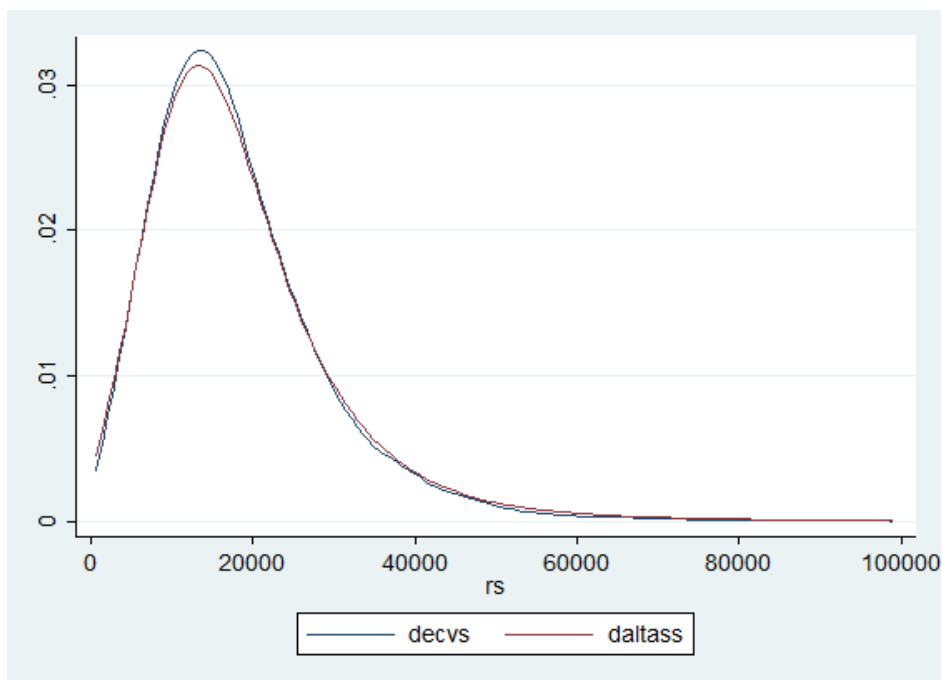


Se trata de una aproximación heurística cuyos resultados dependen del valor de un parámetro conocido como λ . En aplicaciones con series temporales y datos anuales es usual elegir un valor de lambda igual a 100. Esta ha sido también nuestra elección, teniendo en cuenta el número de tramos de renta disponibles. Probando valores alternativos de lambda los resultados son relativamente estables. Frente al empleo de otros procedimientos alternativos de suavización, como los basados en medias móviles,

el filtro empleado tiene la ventaja de que permite obtener todos los valores, incluidos los extremos.

Y, adicionalmente, el empleo de series suavizadas permite comprobar que, sobre todo, en el caso de Comunidades Autónomas, la información que el Atlas ofrece y la información que la ECV ofrece son similares (Gráfico 2).

Gráfico 2. Funciones de densidad de probabilidades suavizadas de la ECV (decvs) y del Atlas (datlass) por nivel de renta (rs)



Análisis del comportamiento de la cola superior de la distribución

Dado que a efectos fiscales nadie está interesado en declarar una renta superior a la que efectivamente percibe, se calcula la ratio de ricos según el Atlas y según la ECV.

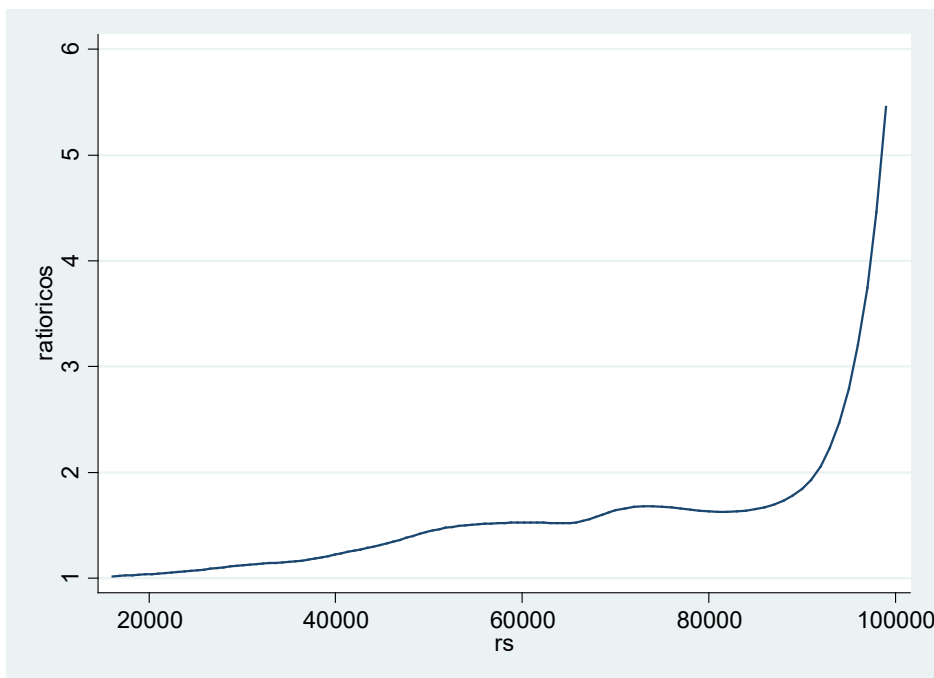
La ratio de ricos se define como:

$$Ratoricos = \frac{1 - patlass}{1 - pecvs} = \frac{\sum_{rs \geq r^u}^{rs=Max} datlass}{\sum_{rs \geq r^u}^{rs=Max} decvs}$$

La ratio de ricos representa el porcentaje de individuos que según el Atlas tienen una renta superior o igual a r^u hasta un máximo de $rs = Max$ dividido por el porcentaje de individuos que según la ECV tienen una renta también superior o igual a r^u hasta alcanzar este mismo máximo.

Los valores elegidos para r^u son el valor de la función de distribución del Atlas en 0.5 y 100,000 € para $rs = Max$. Las frecuencias superiores a 100,000 € son muy reducidas, por lo que toda representatividad quedaría distorsionada (Gráfico 3).

Gráfico 3. Ratio de ricos según Atlas vs ECV por nivel de renta (rs)



El gráfico anterior claramente muestra que la ECV puede infravalorar los valores elevados de la distribución. Es decir, a partir del 50% de la población más pobre, el porcentaje de ricos según el Atlas excede al porcentaje de ricos según la ECV.

En la mayor parte de las CCAA este mismo fenómeno se aprecia con claridad. El patrón es generalizado, en el sentido de aportar más probabilidad para las rentas altas la densidad del Atlas.

Análisis del comportamiento de la cola inferior de la distribución

Tal como se ha dicho en la introducción, en el caso de la cola de rentas bajas el diferencial es menor, y el signo de la comparativa entre ECV y Atlas es distinto según las CCAA. Predomina la captación de rentas bajas por parte de la ECV y no es descartable que las rentas bajas del Atlas tengan en alguna comunidad un sesgo de sobrevaloración.

4.- UNA PROPUESTA DE DISTRIBUCIÓN COMBINADA

Tenemos dos distribuciones de renta, la procedente de la ECV y la procedente del Atlas. Denominamos “p” al valor de la función de distribución procedente de la ECV y “pa” al valor de la función de distribución procedente del Atlas. Denominamos “rs” al nivel superior de renta según los tramos disponibles.

Si a título ilustrativo consideramos un nivel de renta por unidad de consumo de 2,100 € según la ECV un 1.7% de hogares tiene una renta inferior a este nivel mientras que según el Atlas un 2.0% de hogares tiene una renta inferior a este nivel. De estas dos cifras seleccionamos el 1.7%. Es decir, cuando la ratio “pa/p” excede a la unidad, nos quedamos con el valor de “p” como valor más probable.

Lo que ello significa es que, si tenemos dos estimaciones de renta, la declarada en la encuesta o la fiscal, nos quedamos con la más elevada. La declaración fiscal nunca puede exceder a la renta real, pero la renta real si puede exceder a la fiscal. Para niveles de renta bajos es muy posible que este sea el caso.

Consideremos por ejemplo un nivel de renta de 50,400 €. Según la ECV sólo el 1.65% de hogares tienen una renta superior, pero según el Atlas, un 2.54% tienen una renta superior. Dado que este 2.54% procede de registros fiscales, la cifra no observada o real debe ser mayor o igual al 2.54%, motivo por el cual nos quedamos con el valor de “pa”. Como en el caso anterior, lo que ello significa es que, si tenemos dos estimaciones de renta, la declarada en la encuesta o la fiscal, nos quedamos con la más elevada. Para rentas elevadas es muy posible que se de esta situación. De hecho, nunca la renta fiscal puede exceder a la real.

A partir de aquí se define una función de distribución combinada que selecciona “p” cuando la ratio “p/pa” sea superior a la unidad o inferior a la unidad:

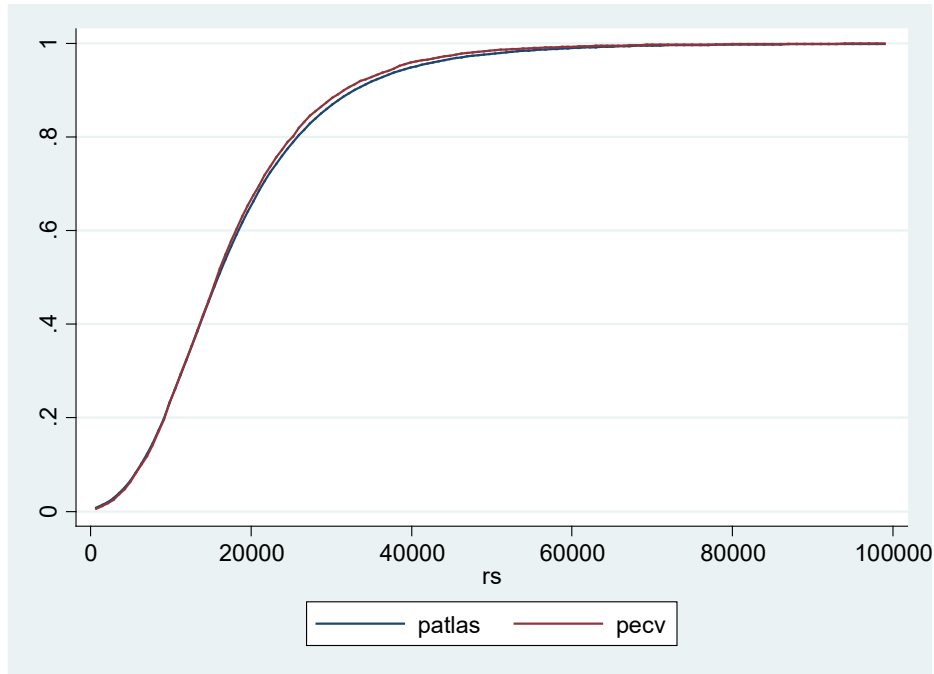
$$\text{Si: } \frac{pa}{p} > 1 \Rightarrow pc = p$$

$$\text{Si: } \frac{pa}{p} < 1 \Rightarrow pc = pa$$

La distribución combinada comportará una renta media algo más elevada que la del Atlas o que la de la ECV, pero entendemos que ello debe ser así en la medida en que el Atlas infravalore las rentas bajas a consecuencia de la economía sumergida y que la ECV infravalore las rentas muy elevadas como los datos demuestran.

Una vez se dispone de los valores de la función de distribución combinada es inmediato obtener la distribución completa y los valores de los indicadores que se consideren relevantes.

Gráfico 4. Funciones de distribución según la ECV (pecv) y el Atlas (patlas), rentas (rs) inferiores a 100,000 €



Para rentas inferiores a rs^j , la ECV nos dice que hay *pecv* observaciones (expresadas en tanto por uno) y el Atlas nos dice que hay *patlas* observaciones. De estos dos valores seleccionamos siempre el más reducido. Es decir, implícitamente estamos ordenando las observaciones de menor a mayor renta de acuerdo con ambas distribuciones y nos quedamos con la proporción más reducida de unidades.

Por tanto, lo que ello implica es que, para los distintos niveles de renta contemplados, al considerar la envolvente, siempre estamos seleccionando una distribución para la cual el nivel medio de renta es el más elevado del que se deriva de cada una de las dos distribuciones individuales.

La fusión de densidad de probabilidades de la envolvente se obtiene como la pendiente de la función de distribución de la ECV cuando es la ECV la que domina, o como la pendiente del Atlas cuando es el Atlas el que domina.

Conceptualmente se tiene:

$$f(r) = \frac{dP(r)}{dr}$$

$$dP(r) = f(r) \cdot dr$$

donde $P(r)$ es el valor de la función de distribución y $f(r)$ es el valor de la función de densidad de probabilidades.

Dado que

$$P(r^0) = \int_{-\infty}^{r^0} f(r) \cdot dr$$

y que $P(\infty) = 1$, debe verificarse que:

$$\int_{-\infty}^{\infty} f(r) dr = 1$$

En nuestro caso, las rentas están acotadas entre cero y un límite superior r^{max} , por lo que la expresión previa adopta la forma:

$$\int_0^{r^{max}} f(r) dr = 1$$

La estimación con datos muestrales de la función de densidad de probabilidades opera con:

$$\widehat{f(r)} = k \cdot \frac{\Delta P(r)}{\Delta r}$$

donde k es un factor de proporcionalidad dado que los datos disponibles no reflejan variaciones infinitesimales sino de tipo discreto.

A efectos de estimar este factor de proporcionalidad y en consonancia con la expresión:

$$\int_0^{r^{max}} f(r) dr = 1$$

se opera de la siguiente forma:

$$\widetilde{f(r)} = \tilde{k} \left(\frac{\Delta P(r)}{\Delta r} \right)$$

$$\sum_{r=0}^{r^{max}} (\widetilde{f(r)}) = \tilde{k} \sum_{r=0}^{r^{max}} \left(\frac{\Delta P(r)}{\Delta r} \right) = 1$$

$$\tilde{k} = \frac{1}{\sum_{r=0}^{r^{max}} \left(\frac{\Delta P(r)}{\Delta r} \right)}$$

En consecuencia:

$$\widetilde{f(r)} = \frac{\left(\frac{\Delta P(r)}{\Delta r} \right)}{\sum_{r=0}^{r^{max}} \left(\frac{\Delta P(r)}{\Delta r} \right)}$$

y esta es la justificación de la restricción que se impone al exigir que la suma de pendientes sea unitaria a efectos de estimar las funciones de densidad de probabilidades a partir de ratios entre magnitudes discretas.

Gráfico 5. Funciones de distribución según la ECV (pecv) y el Atlas (patlas), rentas (rs) entre 0 y 25,000 €

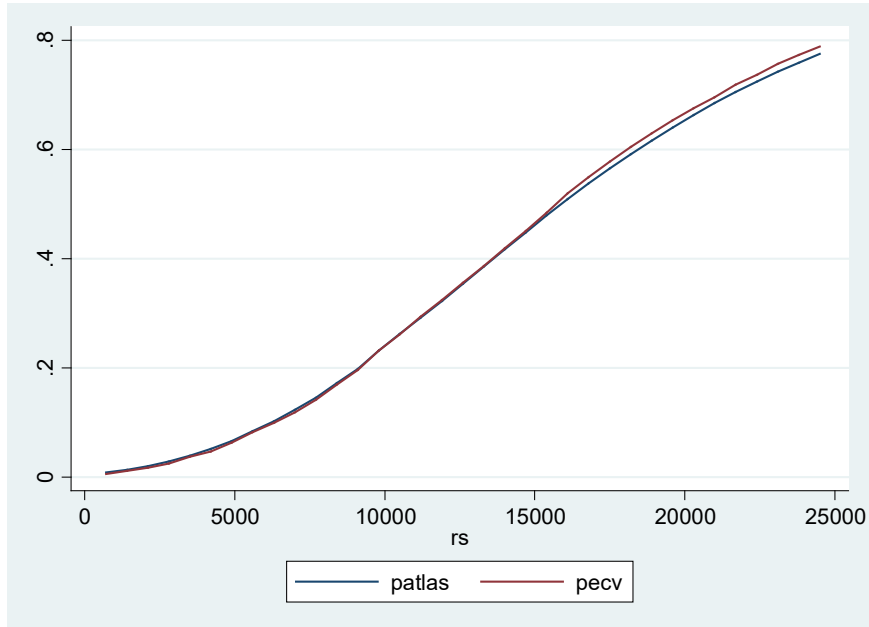


Gráfico 6. Funciones de distribución según la ECV (pecv) y el Atlas (patlas), rentas (rs) entre 25,000 y 50,000 €

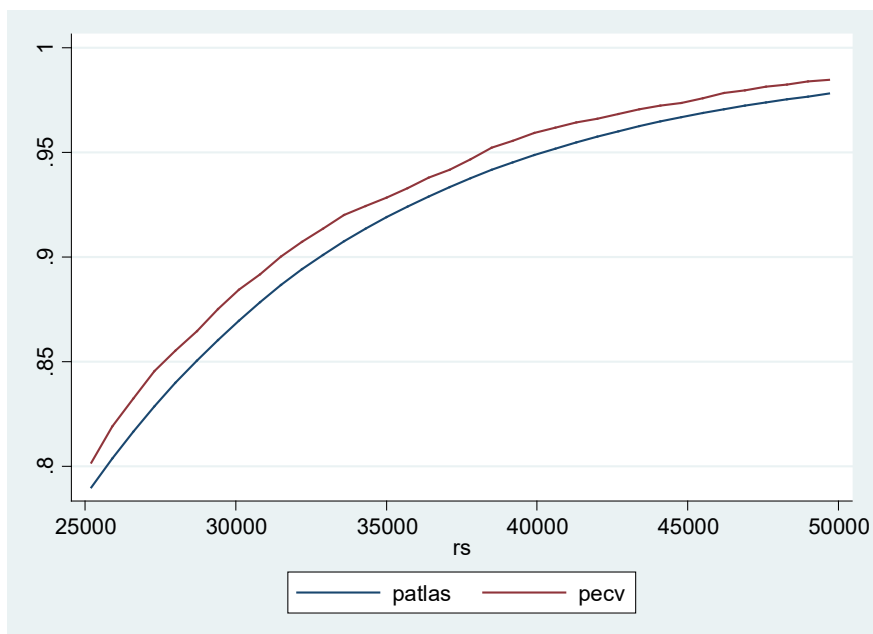


Gráfico 7. Funciones de distribución según la ECV (pecv) y el Atlas (patlas), rentas (rs) entre 50,000 y 75,000 €

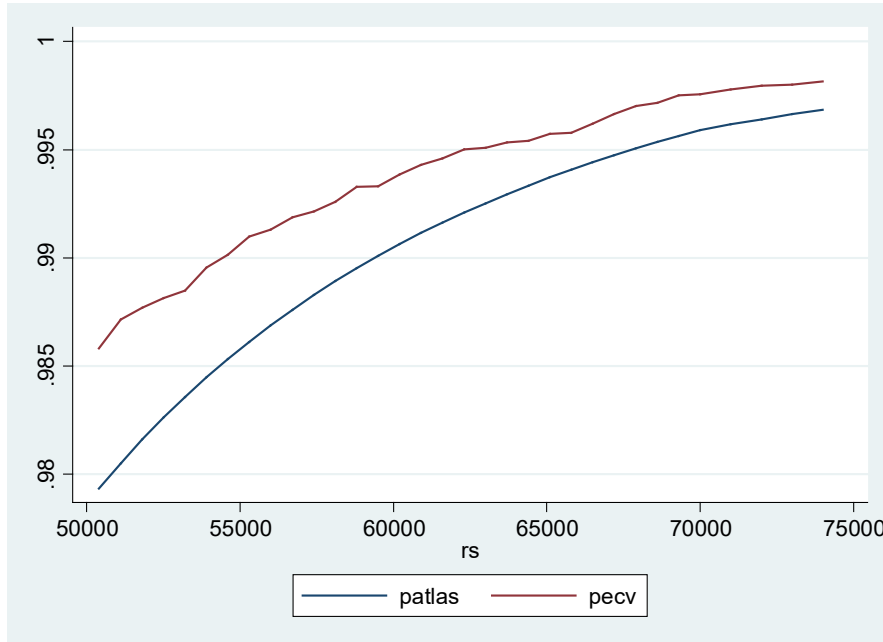


Gráfico 8. Funciones de distribución según la ECV (pecv) y el Atlas (patlas), rentas (rs) entre 75,000 y 100,000 €

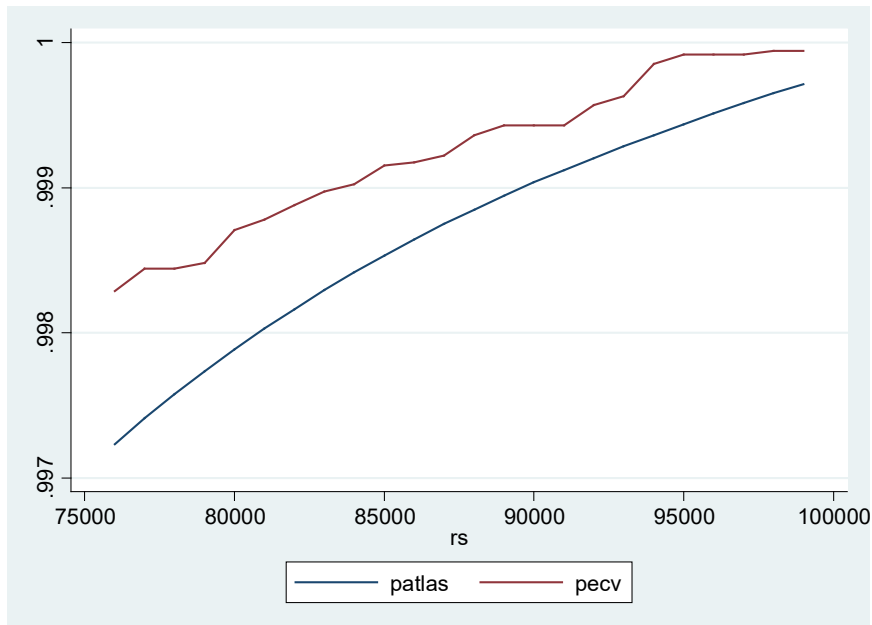


Gráfico 9. Funciones de densidad según la ECV (decv) y el Atlas (datlas) por nivel de renta (rs)

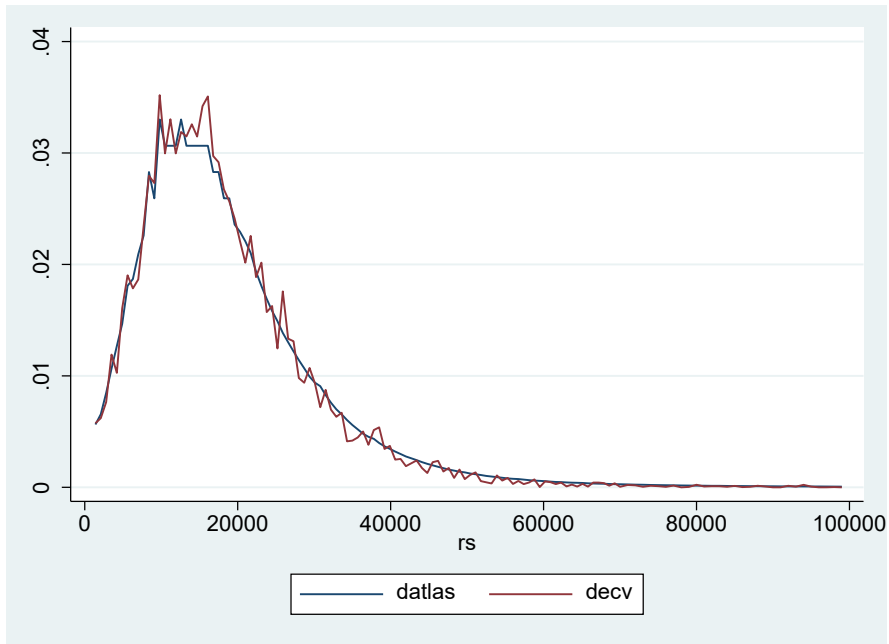


Gráfico 10. Funciones de densidad suavizadas según la ECV (decv) y el Atlas (datlas) por nivel de renta (rs)

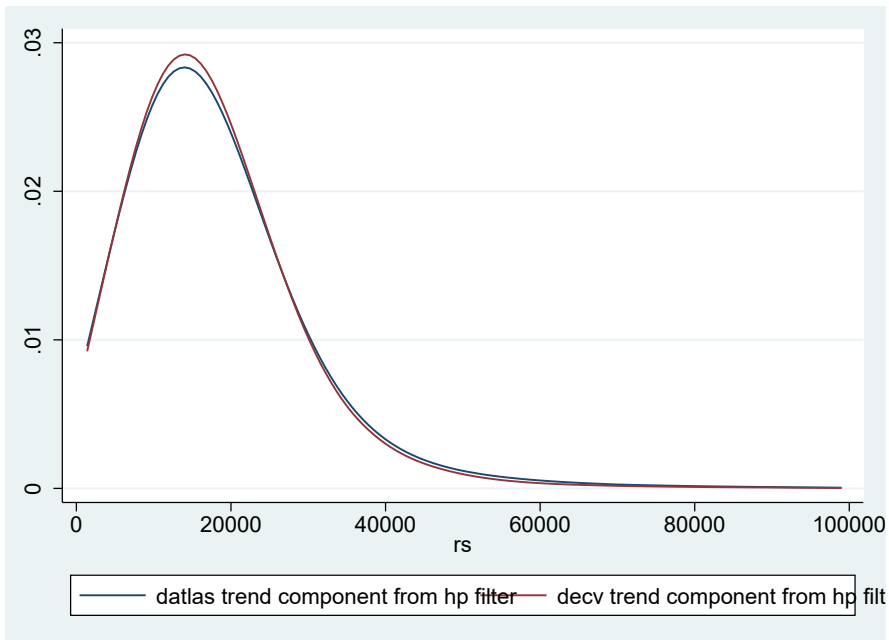
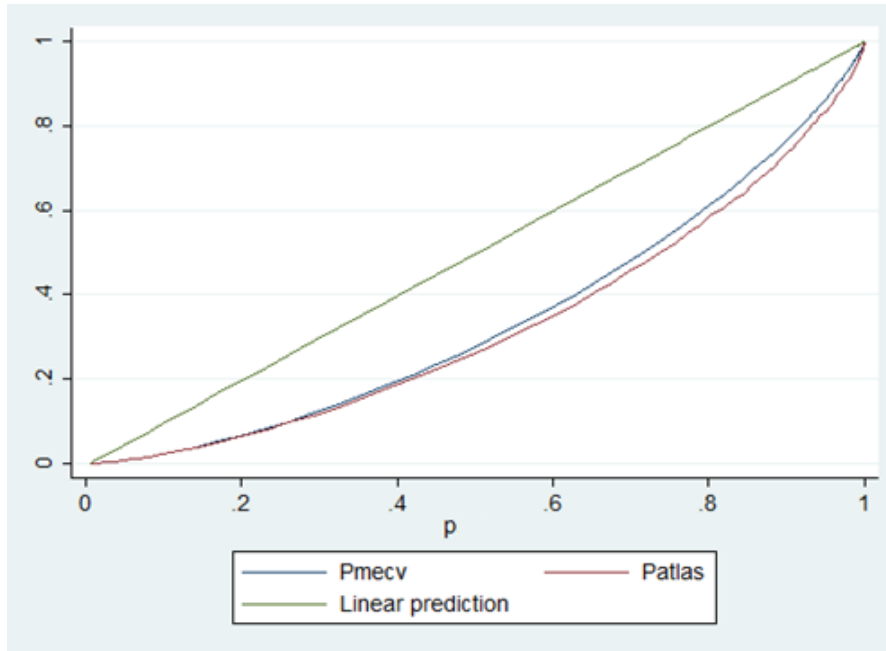


Gráfico 11. Curvas de Lorenz de ECV int (Pmecv) y Atlas (Patlas)



5.- UNA MEDIDA DEL GRADO DE SIMILITUD ENTRE DISTRIBUCIONES

Es interesante estudiar la similitud entre las funciones de densidad originales, ECV y Atlas, con la propuesta combinada de la envolvente.

El grado de discrepancia entre la envolvente y la ECV se define como:

$$\text{Discrepancia ECV vs envolvente: } \sum_{i=1}^N (p_i - pc_i)$$

Para el Atlas se tiene:

$$\text{Discrepancia Atlas vs envolvente: } \sum_{i=1}^N (pa_i - pc_i)$$

La discrepancia total de la envolvente frente a los dos componentes es:

$$\text{Discrepancia total: } \sum_{i=1}^N (p_i - pc_i) + \sum_{i=1}^N (pa_i - pc_i)$$

Igualando a 100 el total, puede calcularse el grado de discrepancia de cada distribución aislada frente a la envolvente. El complemento a 100 nos ofrece el grado de similitud de cada distribución aislada frente a la envolvente tal como la siguiente tabla muestra:

Tabla 2. Grado de similitud respecto a la distribución envolvente propuesta, España (territorio común)

	Diferencia respecto la envolvente	Porcentaje con relación al total diferencias	Grado de similitud
Las diferencias de la envolvente frente a la ECV suman	0.74	85.79	14.21
Las diferencias de la envolvente frente al Atlas suman	0.12	14.21	85.79
Total, diferencias	0.86		

En la Tabla 3 se complementan los resultados anteriores con los relativos a las diferentes CCAA. El análisis de dichos resultados permite destacar:

En el conjunto de España (territorio común) la envolvente es mucho más parecida a Atlas (85,79%) que a ECV (14,21%).

La similitud entre envolvente y Atlas es muy asimétrica porque como muestran las columnas en las que se segmenta el recorrido de la variable renta por unidad de consumo, en el tramo de las rentas bajas hasta la mediana la envolvente se parece más

a la ECV y, en cambio, de la renta mediana al extremo de las rentas altas la envolvente se parece más al Atlas.

Una segunda asimetría se encuentra en la comparativa cuando se aplica a las distintas CCAA. Vemos que la similitud está a la mitad entre CCAA en las que envolvente y ECV se parecen más y el caso contrario. Para el conjunto predomina la similitud entre envolvente y Atlas por el peso demográfico de las CCAA con este perfil.

En las CCAA de mayor renta la ECV capta más rentas bajas y el Atlas capta más rentas altas, Madrid y Catalunya son los ejemplos más claros. En cambio, en las CCAA de renta menor en general siempre capta más renta el Atlas que la ECV, razón por la cual en todos los tramos la envolvente es más parecida al Atlas que a la ECV.

Tabla 3. Similitud de la envolvente con ECV (complemento a 100 con Atlas)

CCAA	Toda la muestra	pc<.25	pc>.25 y pc<.50	pc>.50 y pc<.75	pc>.75
Canarias	0,68	0,00	18,11	0,00	0,13
Murcia, Región de	1,20	24,31	0,00	0,00	0,00
Galicia	1,98	29,33	0,00	0,00	0,33
Andalucía	2,22	8,63	0,00	0,00	3,67
Castilla - La Mancha	2,46	11,37	0,00	0,00	2,34
Asturias	3,08	15,79	0,00	0,00	2,08
Extremadura	5,40	69,23	0,00	0,00	1,32
Baleares, Illes	6,41	98,09	41,07	0,00	0,00
España territorio comun	14,21	100	100	0	0
Comunitat Valenciana	17,14	90,27	34,03	40,76	0,00
Castilla y León	19,45	73,01	94,94	11,37	0,00
Comunidad de Madrid	41,54	100,00	100,00	75,02	1,55
Aragón	43,14	85,12	99,42	98,25	3,50
Cantabria	51,88	100,00	1,52	0,00	44,13
Catalunya	68,14	100,00	100,00	99,08	1,05
Rioja, La	81,61	100,00	100,00	100,00	27,95

6.- COMPARATIVA DE ESTADÍSTICOS

Se presentan resultados para los estadísticos más característicos de estas distribuciones. Se compara ECV, Atlas y el modelo combinado propuesto.

El Riesgo de Pobreza

Los resultados para la pobreza según ECV, Atlas y la envolvente, original y suavizada, aparecen en la Tabla 4. Además, en las dos últimas columnas aparecen los resultados de la pobreza aplicando los datos de Paridad de Poder de Compra (PPA), es decir, si se tiene en cuenta el distinto coste de la vida de las CCAA.

Ya se ha visto que la comparativa entre ECV y Atlas tiene para las CCAA distinto signo y que, en general, las CCAA de menor renta tienen mayor pobreza según ECV y las de mayor renta tienen mayor pobreza según Atlas. Como se aprecia en la Tabla 3, para Extremadura, Andalucía, Asturias y Castilla La Mancha la pobreza era superior en ECV que en Atlas. La envolvente corrige los dos datos, aunque es superior la corrección del dato de ECV que de Atlas.

Por ejemplo, para Extremadura corrige a la baja el riesgo de pobreza de ECV en -3,2 puntos y aumenta el resultado del Atlas en menos de un punto (0,9). El mismo tipo de corrección encontramos en Andalucía, Asturias y Castilla La Mancha.

Por su parte, las CCAA para las que el Atlas tiene mayor pobreza que en la ECV fueron La Rioja, Cantabria, Catalunya, Baleares y Madrid. En este caso, la envolvente corrige las dos fuentes bajando la pobreza, pero de forma más intensa en Atlas y más suave en ECV. El ejemplo de La Rioja es una corrección de la pobreza de Atlas en -3,1 puntos y de -0,6 en ECV. Este tipo de corrección se repite para Catalunya y Madrid.

Por tanto, la envolvente tiende a bajar la pobreza con carácter general, pero sobre todo baja la pobreza de las CCAA con valores más elevados de la misma (Extremadura, Andalucía, Castilla – La Mancha).

Tabla 4. Riesgo de pobreza por CCAA

	POBREZA ECV	POBREZA ATLAS	POBREZA ENV	POBREZA ENV S	POBREZA ENV S PPA	DIF ENV- ECV	DIF ENV- ATLAS	DIF ENV S PPA-ECV	DIF ENV S PPA- ATLAS
Andalucía	32,3	28,5	29,1	28,8	26,3	-3,2	0,6	-6,0	-2,2
Aragón	15,8	15,9	14,9	15,4	14,6	-0,9	-1,0	-1,2	-1,3
Asturias, Principado de	20,4	16,8	17,1	16,9	15,2	-3,3	0,3	-5,2	-1,6
Balears, Illes	17,6	19,4	18,3	18,1	20,0	0,7	-1,1	2,4	0,6
Canarias	28,4	26,6	26,6	26,3	23,1	-1,8	0,0	-5,3	-3,5
Cantabria	15,4	17,2	15,6	14,9	13,7	0,2	-1,6	-1,7	-3,5
Castilla y León	17,9	17,6	17,9	17,0	14,9	0,0	0,3	-3,0	-2,7
Castilla - La Mancha	27,4	24,4	25,3	24,7	18,6	-2,1	0,9	-8,8	-5,8
Catalunya	14,8	16,6	14,2	14,1	16,3	-0,6	-2,4	1,5	-0,3
Comunitat Valenciana	25,1	23,6	23,5	23,0	20,8	-1,6	-0,1	-4,3	-2,8
Extremadura	32,3	28,2	29,1	28,6	21,7	-3,2	0,9	-10,6	-6,5
Galicia	20,2	18,1	18,7	18,4	16,2	-1,5	0,6	-4,0	-1,9
Madrid, Comunidad de	15,2	16,8	14,9	15,1	18,9	-0,3	-1,9	3,7	2,1
Murcia, Región de	27,7	25,2	26,0	25,8	21,3	-1,7	0,8	-6,4	-3,9
Rioja, La	16,1	18,6	15,5	15,4	13,3	-0,6	-3,1	-2,8	-5,3
Ceuta	32,3	33,3	33,0	32,6	32,6	0,7	-0,3	0,3	-0,7
Melilla	30,1	35,3	29,6	30,4	30,4	-0,5	-5,7	0,3	-4,9
PROMEDIO	22,9	22,5	21,7	21,5	19,9	-1,2	-0,7	-3,0	-2,6

En resumen, si se toma como referencia la envolvente, tanto la ECV como el Atlas tienden a sobrevalorar la pobreza, aunque esta sobrevaloración recae sobre distintas CCAA en cada fuente. En ECV se sobrevalora la pobreza de las CCAA de renta baja y en Atlas las de renta alta.

La desigualdad

En la primera parte de la Tabla 4 se presentan los datos del INE para la ECV, el Atlas y la ECV con los datos agrupados de forma idéntica a Atlas (ECVint). Como se dijo anteriormente la conclusión es clara: capta mayor desigualdad el Atlas. Por el estudio de las distribuciones sabemos también la razón: se captan muchas más rentas elevadas en el Atlas que en ECVint, lo que hace aflorar una mayor desigualdad.

Si se asume que la envolvente es una propuesta combinada plausible, sus resultados nos pueden ayudar a valorar los diferentes resultados de ECV y Atlas.

Tabla 5. Índice de Gini por CCAA

	GINI ECV	GINI ECV int (*)	GINI ATLAS	GINI ENV	GINI ENV S	GINI ENV S Ind	DIF ENV- ECVint	DIF ENV- ATLAS	DIF ENV S Ind- ECV
Andalucía	34,3	31,5	33,0	30,9	31,1	33,8	-0,6	-2,1	-0,5
Aragón	28,7	26,1	29,7	26,9	27,1	29,8	0,8	-2,8	1,1
Asturias, Principado de	31,7	28,5	30,3	28,1	28,1	31,4	-0,4	-2,2	-0,3
Baleares, Illes	28,5	25,4	32,8	29,7	29,8	33,4	4,3	-3,1	4,9
Canarias	35,0	31,9	33,8	31,3	31,5	34,6	-0,5	-2,5	-0,4
Cantabria	29,8	27,2	30,1	26,9	27,0	29,6	-0,3	-3,2	-0,2
Castilla y León	29,5	27,1	30,2	27,6	27,7	30,1	0,5	-2,6	0,6
Castilla - La Mancha	31,7	28,8	31,2	29,3	29,5	32,5	0,6	-1,9	0,8
Catalunya	30,9	28,1	32,3	28,5	28,5	31,3	0,3	-3,8	0,4
Comunitat Valenciana	32,2	29,4	32,6	29,9	30,0	32,9	0,5	-2,7	0,7
Extremadura	30,5	27,1	31,1	28,5	28,8	32,4	1,5	-2,6	1,9
Galicia	29,2	26,0	29,7	27,6	27,7	31,1	1,6	-2,1	1,9
Madrid, Comunidad de	33,4	30,4	35,9	31,3	31,4	34,5	0,9	-4,6	1,1
Murcia, Región de	30,1	27,3	31,7	29,9	30,1	33,1	2,6	-1,8	3,0
Rioja, La	29,3	26,6	30,7	26,9	27,0	29,8	0,3	-3,8	0,5
Ceuta	39,7	37,0	42,3	39,6	39,8	42,7	2,6	-2,7	3,0
Melilla	41,2	38,2	42,4	38,2	38,6	41,6	0,0	-4,2	0,4
PROMEDIO	32,1	29,2	32,9	30,1	30,2	33,2	0,9	-2,9	1,1

Fuente: INE y elaboración propia

Se presenta la envolvente y la envolvente suavizada. Se constata que en términos de Gini los resultados son prácticamente idénticos.

La envolvente muestra unos Gini entre los de ECVint y Atlas. Por la magnitud de las diferencias la conclusión es que ECVint infravalora ligeramente la desigualdad, pero el Atlas la sobrevalora, incluso con mayor intensidad que la infravaloración de ECVint.

Para poder comparar la envolvente (obtenida con datos agrupados en intervalos) con la ECV publicada por el INE (Gini con datos individuales) se ha aplicado la ratio de diferencial entre ECV y ECVint a la envolvente. El resultado aparece en la columna GINI ENV S Ind (datos individuales). En la última columna se presentan las diferencias entre nuestra estimación y la publicada por el INE. Se confirmaría que, en términos generales, hay una ligera infravaloración de la desigualdad de la ECV, pero mucho menos que la

indicada por el Atlas. Esto se puede apreciar tanto mirando los resultados de las CCAA como el promedio de todas ellas.

La explicación de estos resultados puede ser la siguiente. El Atlas no recoge bien las rentas bajas y sí las rentas más altas, razón por la cual sobreestima la desigualdad. En cambio, la ECV no recoge bien las rentas altas y por ello subestima la desigualdad. Ahora bien, como el peso de las rentas bajas no captadas por Atlas es superior al peso de las rentas altas no captadas por ECV, el resultado final es que la envolvente, como distribución más plausible, está más cerca de la desigualdad de ECV que de Atlas.

En resumen: la desigualdad captada por la envolvente está algo por encima de los resultados de la ECV y por debajo de los resultados del Atlas, pero más cerca de la ECV que del Atlas.

7.- CONCLUSIONES

Tal como se ha dicho en la introducción, el punto de partida de este trabajo ha sido la discrepancia entre los indicadores de pobreza y desigualdad de dos operaciones del INE: la Encuesta de Condiciones de Vida (ECV) y el Atlas de Distribución de la Renta de los Hogares.

El desarrollo del trabajo puede presentarse en torno a la respuesta a cuatro preguntas. Veamos las preguntas y las respuestas que se han alcanzado.

1. ¿Son significativas las diferencias en los indicadores de pobreza y desigualdad para las CCAA de ECV y Atlas?

Efectivamente, hay diferencias significativas, especialmente en la medición de la desigualdad, tanto por la magnitud de la diferencia como por su patrón sistemático. En términos de los valores de Gini, el Atlas capta más desigualdad en todas las CCAA que

la ECV, cuando se hace una comparativa homogénea, es decir, cuando para ambas fuentes se calcula el Gini con el mismo agrupamiento de datos.

En el caso de la pobreza también se detectan diferencias, aunque son menores y con un patrón menos claro, ya que aparecen diferencias de signo contrario para distintas CCAA. Sin embargo, predominan los casos en los que se capta más pobreza en la ECV. En general, se capta mayor pobreza en ECV en las CCAA de rentas bajas y se capta mayor pobreza en Atlas para las CCAA de rentas altas.

2. Si las diferencias son significativas, y a la vista de las distribuciones de ambas fuentes, ¿cuál es la causa de estas diferencias?

A la vista de las distribuciones de la renta por unidad de consumo de las dos fuentes, la razón de las diferencias en los indicadores reside en un distinto comportamiento en las colas de ambas distribuciones. El Atlas capta más rentas altas. Por su parte, en la cola de las rentas bajas hay menor diferencia, pero en la mayor parte de las CCAA se captan más rentas bajas en las ECV que en el Atlas.

3. ¿Qué resultados puede valorarse como más fiables o realistas?

La respuesta se basa en un argumento económico: no hay incentivos para inflar las rentas, especialmente en las declaraciones de hacienda. Por tanto, en la ECV puede existir una infracobertura de las rentas altas y, en cambio, puede estar captando correctamente rentas bajas de la economía informal. Por su parte el Atlas puede ser más fiable en rentas altas, y, en cambio, puede ser menos fiable en las bajas, ya que para estas rentas sí puede existir ese incentivo (incluso para presentar rentas negativas). Por tanto, para las rentas altas puede ser más fiable el Atlas y para las bajas la ECV. En resumen, dado que la información del Atlas procede de las retenciones fiscales, todas las rentas procedentes de la economía sumergida quedarán excluidas en el Atlas, pero no necesariamente en la ECV.

4. ¿Se puede definir una distribución que recoja lo mejor de cada fuente?

Sobre la base de la anterior valoración, se puede definir una distribución que combina ambas fuentes a través de la función envolvente.

Si se acepta la envolvente como distribución de referencia la conclusión es que, como muestra el Gráfico 12, la ECV tiende a sobrevalorar parcialmente la pobreza en las CCAA de menor nivel de renta y mayor pobreza (Andalucía, Extremadura) y el Atlas tiende a sobrevalorar parcialmente la pobreza, en las CCAA de mayor nivel de renta y menor pobreza (Catalunya o Madrid). En el gráfico, la línea que representa una tasa de riesgo de pobreza del 20% separa las CCAA en las que la pobreza según el Atlas es mayor, de las CCAA en que la pobreza según la ECV es mayor. En relación con la desigualdad, tal como muestra el Gráfico 13, la ECV tiende a infravalorar (ligeramente) la desigualdad y el Atlas tiende a sobrevalorarla, de forma más significativa y con carácter general.

Gráfico 12: La pobreza según ECV, Atlas y Envolvente. 2020. CCAA

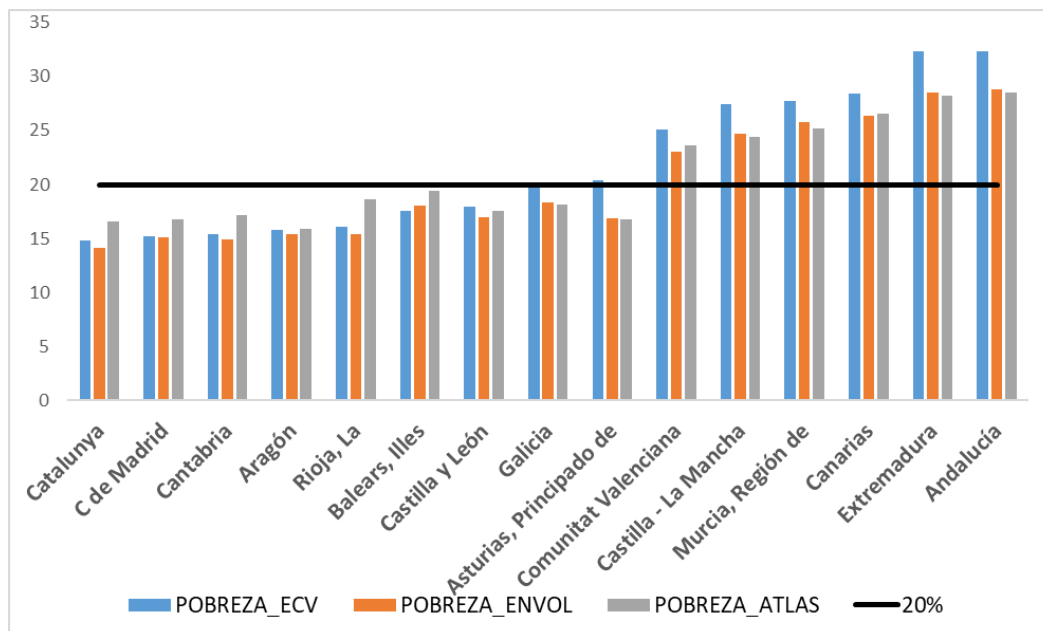


Gráfico 13: La desigualdad según ECV, Atlas y Envolvente. 2020. CCAA

