



Trabajo Fin de Máster

Dos análisis sobre la función redistributiva del sector público en España: diferencias territoriales y capacidad de reacción ante la pandemia COVID-19

Two analyses of the redistributive function of the public sector in Spain: territorial differences and responsiveness to the COVID-19 pandemic

Autor

Pablo Padilla Borrajo

Director

Dr. Jorge Bielsa Callau

Información

Autor: Pablo Padilla Borrado

Director: Jorge Bielsa Callau

Título del trabajo: Dos análisis sobre la función redistributiva del sector público en España: diferencias territoriales y capacidad de reacción ante la pandemia COVID-19

Title of the work: Two analyses of the redistributive function of the public sector in Spain: territorial differences and responsiveness to the COVID-19 pandemic

Titulación: Máster Universitario en Economía

Resumen:

La desigualdad de renta y riqueza es un problema creciente sobre el que el sector público tiene capacidad de actuar a través de impuestos y transferencias. El objetivo de este trabajo es estudiar empíricamente el impacto de esa acción del sector público español, poniendo énfasis en distinguir qué herramientas contribuyen más a esa función redistributiva al tiempo que testeamos si es homogénea a lo largo del territorio y si tuvo la flexibilidad que exigía la pandemia del COVID-19.

Para ello, se toman los datos procedentes de la Encuesta de Condiciones de Vida y se selecciona de entre las diferentes metodologías propuestas en la literatura aquella con las características deseables para el estudio. Los indicadores utilizados son, fundamentalmente, el índice de Gini y el índice de Reynolds-Smolensky. Gracias a la granularidad espacial y a los microdatos de la encuesta, los análisis de los valores y la evolución de esos indicadores se pueden realizar teniendo en cuenta tanto la distribución geográfica como por tramos de renta de impuestos y transferencias.

Los resultados reflejan que el efecto redistributivo del sector público en España se debe principalmente a las transferencias monetarias. Asimismo, se detecta cierta disparidad entre las distintas comunidades autónomas, siendo en las más envejecidas donde mayor peso tienen las transferencias debido, tal y como se refleja en la literatura, al fuerte impacto redistributivo de las pensiones.

En relación con la pandemia y contrariamente a la intuición, el tramo de población que mayor incremento de renta disponible tuvo entre 2019 y 2020 fue el compuesto por los individuos situados entre los percentiles 91 y 99 de la distribución. Adicionalmente, se

observa que la desigualdad aumentó en términos generales durante la pandemia, si bien este aumento estuvo mitigado por la acción del sector público gracias al aumento de su esfuerzo redistributivo realizado en ese año.

Abstract:

Income and wealth inequality is a growing problem on which the public sector has the capacity to act through taxes and transfers. The aim of this paper is to empirically study the impact of this action by the Spanish public sector, with emphasis on distinguishing which tools contribute most to this redistributive function, while testing whether it is homogeneous across the territory and whether it had the flexibility required by the COVID-19 pandemic.

To this end, we take data from the Encuesta de Condiciones de Vida and select from among the different methodologies proposed in the literature the one with the desirable characteristics for the study. The indicators used are, fundamentally, the Gini index and the Reynolds-Smolensky index. Thanks to the spatial granularity and the micro-data of the survey, the analysis of the values and the evolution of these indicators can be conducted considering both the geographical distribution and the income brackets of taxes and transfers.

The results show that the redistributive effect of the public sector in Spain is mainly due to monetary transfers. Furthermore, we detect a certain disparity between the different autonomous communities, with the transfers having the greatest weight in the most aged regions, due, as reflected in the literature, to the strong redistributive impact of pensions.

In relation to the pandemic and counter-intuitively, the segment of the population with the largest increase in disposable income between 2019 and 2020 was composed of individuals located between the 91st and 99th percentiles of the distribution. We also observe that inequality increased overall during the pandemic, although the public sector's action mitigated this increase thanks to its increased redistributive effort in that year.

Índice

1.-. Introducción	4
2.-. Medición de la desigualdad de la renta	7
2.1.-. Índice de Gini	7
2.2.-. Índice de Reynolds-Smolensky	9
3.- Metodología y datos	12
4.- Resultados	22
4.1.-. Efectos del sector público en las distintas regiones españolas.....	22
4.2.-. Efectos del sector público en la pandemia Covid-19	28
5.- Conclusiones	34
6.- Referencias bibliográficas	36
Anexos	41
Anexo I. Código empleado	41

1.- Introducción

La cuestión de la desigualdad es un tema de estudio fundamental dentro de las ciencias sociales y se ha abordado desde perspectivas muy distintas. A su vez, se trata de un problema de primer orden tal y como pone de manifiesto Naciones Unidas al incluirlo como el décimo objetivo de la Agenda 2030 y afirmar que “la desigualdad amenaza el desarrollo social y económico en el largo plazo, frena la reducción de la pobreza y destruye el sentido de realización y autoestima de las personas” (Naciones Unidas, 2024).

Para justificar por qué no es deseable la desigualdad, desde una perspectiva neocontractualista, Rawls (1971) plantea un criterio tipo Maximin a través de la Teoría de la Justicia Social. Según este planteamiento, en un estado original donde existe el velo de la ignorancia (se desconoce cuál va a ser el lugar que uno va a ocupar en la escala de renta), los individuos preferirían una situación donde se produjera una redistribución hacia aquellos de menor renta ante la posibilidad de ser ellos los que se situaran en esa situación.

Por diferentes motivos, en la última década se ha publicado abundante literatura sobre desigualdad desde la ciencia económica. Sin ánimo de ser exhaustivos, podríamos citar algunas aportaciones relevantes respecto al asunto que trata este estudio. Por ejemplo, algunos autores como Stiglitz (2012) han expuesto el alto coste que puede tener para las sociedades la presencia prolongada de situaciones de desigualdad. En este mismo sentido, Ostry et al. (2014) señalan que la desigualdad puede debilitar el avance social y generar inestabilidad política y económica que a su vez desincentive la inversión y disminuya el consenso social necesario para mitigar los efectos de posibles shocks macroeconómicos negativos, limitando en última instancia el ritmo y la sostenibilidad del crecimiento económico. Por último, Acemoğlu et al. (2015) se centran en las consecuencias sobre la calidad de las instituciones y el sistema democrático del aumento de la desigualdad. Esta derivada tiene a su vez importantes implicaciones económicas ya que, según estos autores, la fragmentación social acabaría provocando conflictos que a su vez generan ineficiencias en el normal desempeño de la economía.

Asimismo, incluso desde una perspectiva puramente individualista, autores como Friedman (1962) han llegado a la conclusión de que la reducción de la desigualdad resulta positiva también para los individuos que no se ven beneficiados directamente de la

redistribución y, por tanto, las políticas redistributivas son útiles para el conjunto de la sociedad.

En definitiva, la desigualdad es una problemática extremadamente importante para la cual se asume que el sector público tiene una importante misión compensatoria a través de mecanismos de redistribución de la asignación de las rentas privadas que se derivan del mercado. No en vano, esta función de redistribución de la renta es, de acuerdo con Musgrave (1939), una de las tres funciones del sector público. Por tanto, queda plenamente justificado el estudio de la acción del sector público en su función de redistribución de la renta con la finalidad de reducir la desigualdad.

En la literatura se ha evidenciado cómo los impuestos y las transferencias son capaces de reducir la desigualdad observada en la renta de mercado en comparación con la renta disponible, si bien, se ha visto que estos efectos varían en función del país (Joumard et al., 2013). Para el caso español, existen trabajos que, además, demuestran que este efecto es mayor para las transferencias que para los impuestos (López Laborda et al., 2023).

Adicionalmente, en el caso concreto español, siendo un país tan descentralizado, parece razonable estudiar si el sector público está teniendo un efecto similar reduciendo las desigualdades en todas las comunidades autónomas, así como si los instrumentos de los que se dispone están teniendo efectos homogéneos en el territorio.

En tercer lugar, disponemos en España de un experimento natural puntual de esfuerzo redistributivo, la pandemia COVID-19. Este hecho no solo cambió durante unos meses el estilo de vida de todos los ciudadanos, sino que supuso además un reto por parte de los estados por la necesidad de adoptar medidas redistributivas nunca vistas.

En este sentido, se ha observado que las pandemias tradicionalmente afectan más a las personas con menos recursos (Wade, 2020) y se han realizado algunos trabajos, especialmente al comienzo de la pandemia, buscando anticipar los resultados que se iban a dar en términos de desigualdad, por ejemplo, en base a los movimientos de las tarjetas bancarias de los ciudadanos (Aspachs et al., 2021).

No obstante, en este trabajo nos queremos centrar en la evolución de la desigualdad entre 2019 y 2020, tanto antes como después de la intervención pública, analizando si esta ha tenido un impacto significativo en su reducción.

En definitiva, el objetivo principal de este trabajo es estudiar el impacto en la desigualdad del sector público poniendo especial énfasis en tres aspectos: en primer lugar, analizar si son las transferencias o los impuestos los que más contribuyen a este papel redistributivo del Estado; en segundo lugar, explotar la información geográfica para testear la homogeneidad de la redistribución en el territorio; y, por último, estudiar, por comparación entre 2019 y 2020, el esfuerzo redistributivo que se realizó durante la pandemia.

Para alcanzar los objetivos expuestos anteriormente el trabajo se divide en cinco capítulos, contando cada uno de ellos con un objetivo específico. En el primer capítulo, esta introducción, tratamos de presentar el tema de estudio. En el capítulo segundo presentamos alguna de las principales metodologías para medir la desigualdad, sus principales problemas y cómo proponemos solventarlos en este caso concreto. En el tercer capítulo se estudian distintas alternativas que se han realizado hasta la fecha para la obtención de la renta de los individuos a partir de la Encuesta de Condiciones de Vida (ECV) del Instituto Nacional de Estadística (INE) siguiendo la estela de trabajos anteriores como Goerlich (2016, 2018 y 2020) o Bautista (2020). A partir de esa base, se selecciona un conjunto de métodos que, a nuestro juicio, cumplen las características deseables y son operativos.

A continuación, en el cuarto capítulo, se exponen los resultados obtenidos de dos formas complementarias tratando los temas anteriormente mencionados. En primer lugar, se trata de descifrar las diferencias existentes en el papel redistributivo del sector público entre las distintas comunidades autónomas españolas. En segundo lugar, se busca evaluar la labor del sector público como reductor de la desigualdad en la crisis, primero sanitaria y posteriormente, económica, de la COVID-19. Finalmente, en el quinto capítulo se resumen las conclusiones alcanzadas en este trabajo acerca de la función redistributiva del sector público en el caso concreto de estudio.

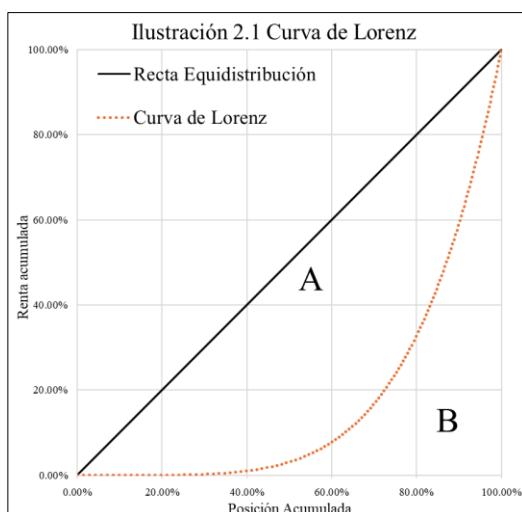
2-. Medición de la desigualdad de la renta

La primera cuestión que se nos plantea es cómo medir la desigualdad de la renta, puesto que este será el punto de partida para todo el trabajo. En la literatura se han propuesto numerosos índices y herramientas para medir la desigualdad de la renta, llegando Coulter (1989) a contabilizar más de 50 índices diferentes. Además, en los últimos tiempos, la investigación a este respecto no ha disminuido su ritmo (McGregor et al., 2019) y, además, las nuevas tecnologías han permitido desarrollar mediciones más específicas, ya sea a través de machine learning aplicado a satélites (Jean et al., 2016) o datos recogidos a través de los teléfonos móviles (Blumenstock et al., 2015).

2.1-. Índice de Gini

Por encima de todos los índices y medidas, el más extendido es el llamado índice de Gini (Gini, 1912 y 1914). Este indicador, además de ser uno de los principales instrumentos en la cuantificación de la desigualdad de la renta, ha sido utilizado en otros ámbitos tan dispares como la astrofísica, la ingeniería o la medicina (Sithiyot y Holasut, 2020). El índice de Gini se puede derivar directamente de la curva de Lorenz (Lorenz, 1905) en la cual se representa en el eje de abscisas la posición acumulada que ocupa cada individuo ordenado de menor a mayor renta y en el eje de ordenadas la renta acumulada normalizada desde el individuo con menor renta hasta el individuo con mayor renta.

Así, el índice de Gini se puede obtener, de acuerdo con Gini (2005), como la ratio entre el área entre la recta que marca la diagonal del cuadrado y la curva de Lorenz (A) y el área que queda por debajo de la recta de equidistribución (A+B). Tal y como se muestra la siguiente ilustración:



Fuente: Elaboración propia.

Por tanto, el índice de Gini puede tomar valores entre 0, donde la igualdad es máxima y todos los individuos tendrían la misma renta y 100, donde un individuo concentraría toda la renta.

Para calcular el índice de Gini se va a seguir la siguiente expresión (Cowell, 2000):

$$100 \times \sum_{i=1}^n k(i)x(i),$$

donde (1)

$$k(i) = \frac{w(i)}{\sum_{i=1}^n w(i)} \left(2 \sum_{j=1}^i w(j) - w(i) - 1 \right)$$

$x(i)$ es la observación correspondiente al i -ésimo valor más pequeño de la muestra y $k(i)$ un factor de ponderación que depende de los pesos de las observaciones $w(i)$.

Esta especificación es, según Cowell (2000), quizá la más ventajosa en términos de computación, razón por la que probablemente haya sido elegida para calcular el índice por el paquete de R utilizado, acid (Sohn, 2016). Cabe destacar que se han verificado los cálculos mediante la réplica de los resultados con otro método de cálculo empleado por el paquete DescTools (Signorell, 2024) obteniendo resultados coincidentes.

No obstante, existen críticas acerca de la utilización del índice puesto que requiere de asunciones acerca de la forma de la función de bienestar social subyacente que nos son deseables (Atkinson, 1970) o porque da demasiada importancia a la parte central de la distribución (Cobham y Sumner, 2013). A pesar de estas críticas, dado que es la medida por antonomasia de desigualdad, el primero de los instrumentos que se van a utilizar en el presente trabajo es el índice de Gini.

Cabe destacar que los valores obtenidos de la renta pueden ser negativos cuando se tienen en cuenta las pérdidas de los autónomos o empresarios. En estos contextos, la curva de Lorenz se sitúa al comienzo por debajo del eje de ordenadas lo que puede generar que el coeficiente de Gini pueda tomar valores mayores que 100 (Hagerbaumer, 1937), algo que puede llegar a dificultar su interpretación.

Para abordar estos problemas, las prácticas más comunes son: o eliminar aquellas observaciones donde la renta sea negativa o bien transformar dicha renta a 0. De hecho, esta última práctica es la recomendada por la OCDE (2015).

En este trabajo, dado que el porcentaje de individuos con renta negativa es reducido, se ha optado por seguir el criterio propuesto por De Battisti et al. (2019) e incluir estas observaciones en la muestra. En este sentido, se ha verificado que no se corre el riesgo de tener índices de Gini superiores a 100 y la distorsión que se produce respecto a otras propuestas metodológicas es ínfima.

Asimismo, se ha tomado la decisión de no acotar los individuos situados en el 1,5% superior e inferior de la distribución como hace el INE (2024) con la finalidad de no alterar los resultados.

2.2-. Índice de Reynolds-Smolensky

Una vez disponemos de una medida acerca de la desigualdad de un territorio, necesitamos una herramienta que nos permita analizar cómo de redistributiva es la acción del sector público.

En este contexto las principales medidas que se suelen emplear son el índice de Kakwani (1977), que mide la progresividad de una intervención pública y el índice de Reynolds-Smolensky (1977), que permite cuantificar la capacidad redistributiva de las intervenciones. En particular, en este trabajo, dado que lo que nos interesa medir es la capacidad redistributiva, se ha decidido emplear el índice de Reynolds-Smolensky.

Este índice se basa en los conceptos de la curva de Lorenz y el índice de Gini. Así, partimos de una distribución inicial de la renta para la cual calculamos un índice de Gini de acuerdo con la ecuación 1. Ahora bien, tras la acción del sector público, se produce una redistribución entre las rentas de los individuos, para estas nuevas rentas calculamos un coeficiente de concentración de acuerdo con la siguiente expresión:

$$100 \times \sum_{i=1}^n k(i)x(i),$$

donde (2)

$$k(i) = \frac{w(i)}{\sum_{i=1}^n w(i)} (2 \sum_{j=1}^i w(j) - w(i) - 1)$$

$x(i)$ es la observación correspondiente al i -ésimo valor más pequeño de la muestra original y $k(i)$ un factor de ponderación que depende de los pesos de las observaciones

$w(i)$. Es decir, tenemos un cálculo equivalente al del índice de Gini mostrado en la ecuación 1 con la salvedad de que la ordenación es la de la distribución inicial.

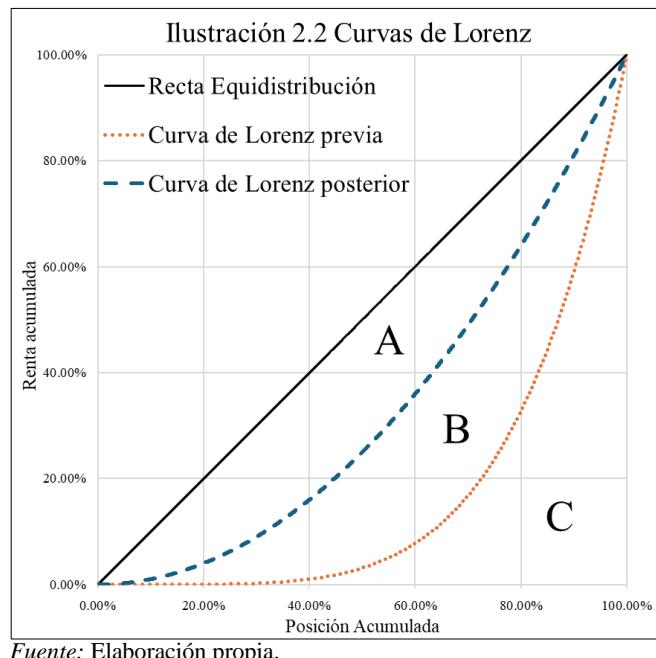
Posteriormente, el índice de Reynolds-Smolensky (Π^{RS}) se obtiene como:

$$\Pi^{RS} = G_{X_1} - C_{X_2} \quad (3)$$

Donde G_{X_1} representa el índice de Gini de la distribución de la renta X_1 y C_{X_2} el índice de concentración de la renta X_2 , es decir, el índice de Gini manteniendo la distribución de la renta que había en X_1 . Este índice puede tomar valores entre -100, donde la capacidad la renta pasaría de distribuirse equitativamente entre todos los individuos a concentrarse en un solo individuo y la acción redistributiva sería mínima, y 100, donde se pasaría de una renta concentrada en un individuo a una distribución igualitaria entre todos los individuos y, por tanto, la acción redistributiva sería máxima.

Cabe mencionar que, en el caso de que no se produjeran reordenaciones, el índice de concentración y el índice de Gini serían iguales ($C_{X_2} = G_{X_2}$). No obstante, como este supuesto es demasiado restrictivo, resulta fundamental mantener la distribución inicial de la renta en el cálculo de los índices de concentración.

El índice de Reynolds-Smolensky se puede obtener a su vez de forma gráfica. Así, tal y como se muestra en la siguiente ilustración, el indicador sería la ratio entre el área entre las curvas de Lorenz antes y después de la intervención (B) y el área que queda por debajo de la recta de equidistribución (A+B+C).



Dado que uno de los objetivos de este trabajo es analizar el efecto redistributivo del sector público y, en particular, cuantificar qué parte de ese efecto viene por las transferencias monetarias y qué parte por los impuestos directos, necesitamos una medida que nos permita llevar a cabo esa misión. Así, partiendo de la ecuación 3, siguiendo el razonamiento expuesto por Bautista (2020) se pueden realizar las siguientes modificaciones:

$$\Pi^{RS} = G_X - C_{X+B-T} = (G_X - C_{X+B}) + (C_{X+B} - C_{X+B-T}) = \pi_B^{RS} + \pi_T^{RS} \quad (4)$$

Donde Π^{RS} es el índice de Reynolds-Smolensky de la acción del sector público, G_X el índice de Gini de la renta de mercado y C_{X+B-T} el índice de concentración de la renta disponible, es decir, el índice de Gini de la renta de mercado X más transferencias monetarias B menos impuestos directos T manteniendo la ordenación observada en la renta de mercado. Así, somos capaces no solo de obtener el índice total, sino también de poder diferenciar el efecto redistributivo tanto de las transferencias monetarias π_B^{RS} como de los impuestos directos π_T^{RS} .

3-. Metodología y datos

El objetivo del presente apartado es presentar la metodología que se va a emplear a lo largo del trabajo así como detallar las asunciones empleadas en el mismo y las fuentes de donde se obtienen los datos.

La fuente principal de datos del trabajo es la Encuesta de Condiciones de Vida (en adelante, ECV) realizada de forma anual por el INE de acuerdo con el Plan Estadístico Nacional por lo que su cumplimentación es obligatoria (Ley 12/1989, de 9 de mayo y RDL 1110/2020, de 15 de diciembre).

Asimismo, cabe destacar que esta encuesta se realiza de acuerdo con los reglamentos comunitarios que regulan la European Survey on Income and Living Conditions (EU-SILC) por lo cual se considera que tiene unos estándares de calidad adecuados para poder realizar estudios con la suficiente profundidad.

La fuente principal de esta encuesta es la realización de entrevistas personales presenciales, telefónicas y por internet. Asimismo, para las variables relativas a los ingresos se combina la información obtenida con las encuestas con otros ficheros en posesión de la administración ya sea de la Agencia Tributaria u otros organismos.

Cabe mencionar que esta encuesta es la fuente de datos de numerosas investigaciones, ya sea en relación con los ingresos de la población o con determinadas variables de pobreza (Ayala, 2016).

Tal y como se ha mencionado previamente en la introducción, uno de los objetivos de este trabajo es obtener una serie homogénea y comparable de ingresos de los hogares de las distintas comunidades autónomas a partir de los componentes de la ECV. A lo largo de la literatura se han propuesto numerosas metodologías para estimar estas variables (Goerlich, 2016, 2018 y 2020; Bautista, 2020; López Laborda et al., 2023 y Araoz, 2024).

En la metodología finalmente elegida se persiguen principalmente dos objetivos. En primer lugar, obtener diversas variables con las características adecuadas para el correcto desarrollo del estudio. Por otra parte, obtener una serie sea homogénea, es decir, que la variable que mida la renta disponible de los hogares coincida con la variable renta disponible de los hogares (*vhRentaa*) incluida en los ficheros de la ECV, ya que esa es la que toma el INE para el cálculo de diversos indicadores.

Para comenzar con el cálculo de variables, podemos definir en primer lugar la renta de mercado, siendo esta la renta con la que el mercado remunera tanto el capital como el trabajo (incluyendo tanto el trabajo por cuenta propia como el trabajo por cuenta ajena) de los individuos del hogar. Se pueden distinguir cuatro componentes de la renta de mercado:

1. Rentas salariales: formadas por el conjunto de renta brutas monetarias o cuasi monetarias (PY010G) así como de las rentas brutas no monetarias (PY021G) de los trabajadores por cuenta ajena.
2. Cotizaciones sociales a cargo del empleador (PY030G).
3. Renta de los trabajadores por cuenta propia: beneficios o pérdidas monetarios brutos de trabajadores por cuenta propia (autónomos), incluyendo derechos de propiedad intelectual. Cabe destacar que esta variable puede tomar valores negativos lo que hará que algunos hogares cuenten con renta negativa.
4. Rentas del capital y la propiedad: formadas por intereses, dividendos y ganancias de inversiones de capital (HY090G), rentas procedentes del alquiler de propiedades o terrenos (HY040G) y rentas percibidas de esquemas privados de pensiones (PY080G).

Así, la renta de mercado queda definida de acuerdo con la nomenclatura de la ECV como:

$$\begin{aligned} \textbf{Renta de Mercado} = & PY010G + PY021G + PY030G + PY050G + \\ & HY090G + HY040G + PY080G \end{aligned} \quad (5)$$

Acerca de los componentes de la renta de mercado cabe hacer algunas consideraciones por la disparidad de criterios existentes para su consideración a lo largo de la literatura. En primer lugar, Goerlich (2016), Bautista (2020) y Araoz (2024) incluyen como variable que mide las rentas brutas no monetarias de los asalariados la variable PY020G, mientras que en Goerlich (2018 y 2020) utiliza la variable PY021G, empleada en este trabajo, dicha variable solo incluye como ingreso no monetario el vehículo de empresa y los costes derivados de su uso.

Por otra parte, los trabajos de Goerlich (2016) y Bautista (2020) incluyen el valor de mercado de los bienes producidos para autoconsumo (PY070G/HY170N), mientras que ni Goerlich (2018 y 2020) ni Araoz (2024) ni este trabajo los tienen en cuenta al

considerar que su imputación a precio de mercado en lugar de a precio de coste puede sobrevalorar la renta de mercado.

Finalmente, hay dos diferencias entre los trabajos de Goerlich (2016, 2018), Bautista (2020) y Araoz (2024) frente a Goerlich (2020) y el presente trabajo. Primeramente, en los cuatro trabajos mencionados se incluye como parte de la Renta de mercado la variable renta bruta percibida por los menores de 16 años (HY110G), mientras que en este y en Goerlich (2020) no. Tal y como señala Goerlich (2020), el fichero que incluye los datos de los adultos de la ECV solo recoge a aquellos miembros del hogar con una edad en el año de referencia mayor o igual a 16 años y el fichero de miembros del hogar no tiene en cuenta información de tipo económico. Por este motivo, no es posible distinguir si esta renta bruta percibida por menores de 16 años se puede considerar rentas del trabajo por cuenta ajena (algo extraño si tenemos en cuenta la edad legal para trabajar en España), prestaciones sociales o rentas del capital. El hecho de no incluirlo como parte de la renta de mercado y sí, como se verá a continuación, como parte de la renta bruta lleva implícito que se estima más adecuado considerar como transferencias sociales porque resulta más razonable.

Asimismo, los cuatro trabajos mencionados previamente no tienen en cuenta como parte de la renta de mercado las cotizaciones sociales a cargo del empleador (PY030G), mientras que tanto en este trabajo como en Goerlich (2020) sí que se incluyen. Aquí resulta especialmente curioso como Goerlich (2016) señala que “se excluyen de las rentas del asalariado las cotizaciones sociales a cargo del empleador (PY030G), ya que aunque representan un coste salarial para el empleador, no forman parte del salario bruto y no suponen una renta directa desde el punto de vista del trabajador”, cuando, el propio Goerlich (2020) menciona “obsérvese que las cotizaciones sociales a cargo del empleador (PY030G) forman parte de la renta de mercado en tanto que se asimilan al salario bruto”. La decisión de incluirlas en el presente trabajo se basa en este último argumento: dado que se busca tener en cuenta todos los ingresos que están recibiendo los individuos, se estima que las cotizaciones sociales pagadas por el empleador se asemejan a un pago diferido de salarios y, por tanto, han de considerarse como renta de mercado.

Una vez se ha definido la renta de mercado, para la obtención de la renta bruta definida como la suma de todas las rentas, tanto públicas como privadas que reciben los individuos y los hogares, hay que agregar a la renta de mercado el conjunto de transferencias. Estas

transferencias las podemos agrupar en cuatro subgrupos, siendo los tres primeros transferencias públicas y el último privadas:

1. Prestaciones por jubilación (PY100G) y supervivencia (PY110G).
2. Prestaciones por desempleo (PY090G).
3. Resto de prestaciones sociales, incluyendo ayuda por familia e hijos (HY050G), ingresos por asistencia social (HY060G), ayuda para vivienda (HY070G), prestaciones por enfermedad (PY120G), prestaciones por invalidez (PY130G), ayuda para estudios (PY140G) y renta bruta percibida por los menores de 16 años (HY110G).
4. Transferencias periódicas monetarias percibidas de otros hogares (HY080G).

En conclusión, la renta bruta se define atendiendo a la nomenclatura de la ECV como:

$$\begin{aligned}
 \textbf{Renta Bruta} = & \text{Renta de Mercado} + PY100G + PY110G + PY090G + \\
 & HY050G + HY060G + HY070G + PY120 + PY130G + PY140G + \\
 & HY110G + HY080G
 \end{aligned} \tag{6}$$

Respecto de esta variable únicamente cabe recalcar la decisión mencionada anteriormente de incluir dentro de la renta bruta y no de la renta de mercado la renta percibida por los menores de 16 años (HY110G).

Por otra parte, llamamos renta disponible del hogar a la renta de la que dispone un hogar para abordar sus decisiones de consumo y ahorro. Para llegar a ella hemos de disminuir de la renta bruta los impuestos directos pagados por los hogares las cotizaciones sociales a cargo del empleador, así como las transferencias monetarias enviadas a otros hogares. Distinguimos, por tanto, tres componentes:

1. Impuestos directos: contando los impuestos directos al patrimonio (HY120G) y los impuestos sobre la renta y cotizaciones sociales a cargo del empleado (HY140G).
2. Cotizaciones sociales a cargo del empleador (PY030G).
3. Transferencias periódicas enviadas a otros hogares (HY130G).

Así, la renta disponible queda, de acuerdo con nomenclatura de la ECV, como:

$$\begin{aligned}
 \textbf{Renta Disponible} = & \text{Renta Bruta} - HY120G - HY140G - PY030G - \\
 & HY130G
 \end{aligned} \tag{7}$$

Recordemos que, tal y como se ha mencionado previamente, tanto en este trabajo como en el de Goerlich (2020) se incluye la variable cotizaciones sociales a cargo del empleador (PY030G) a diferencia de los trabajos citados previamente en los que no se deduce dicha variable de la renta bruta.

Finalmente, de acuerdo con los objetivos del presente trabajo se hace necesaria la construcción de una variable que recoja exclusivamente la renta que obtienen los hogares derivada de su participación en el mercado, tal y como establece Bautista (2020). Así pues, esta variable denominada renta privada se obtiene sumando a la renta de mercado las transferencias recibidas de otros hogares (HY080G) y restando las transferencias emitidas hacia otros (HY130G).

Por tanto, la renta privada la podemos definir como:

$$\textbf{Renta Privada} = \textit{Renta de Mercado} + \textit{HY080G} - \textit{HY130G} \quad (8)$$

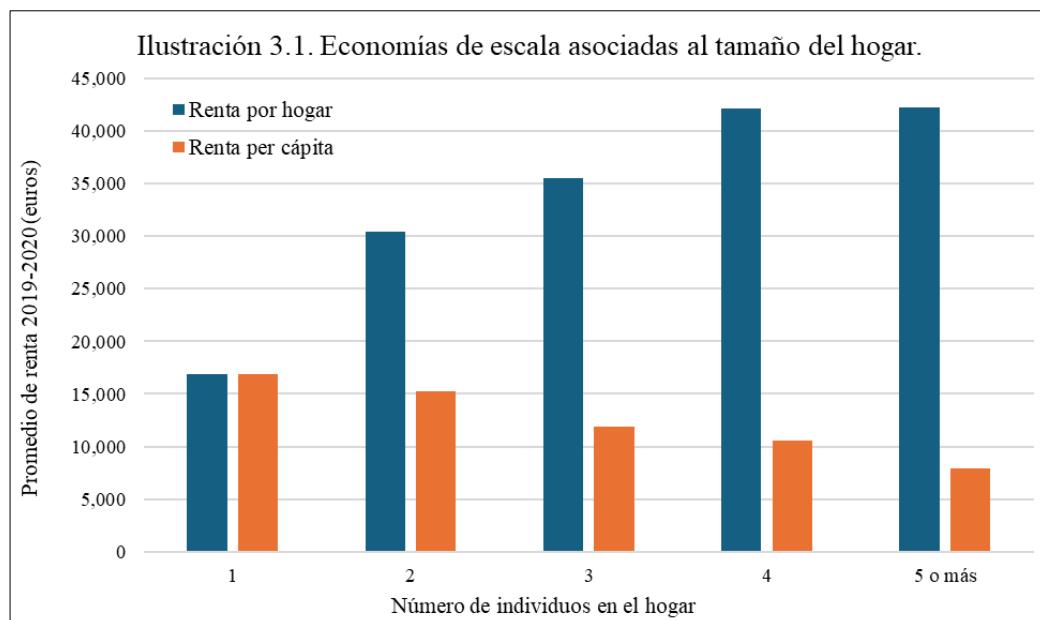
En el caso de que se estuviera ante una economía cerrada, las transferencias recibidas por los hogares serían iguales a las transferencias emitidas ya que, al no haber sector exterior, lo que reciben unos hogares españoles debería ser igual a lo emitido por otros. En el caso particular de España, el total de transferencias recibidas por los hogares en el año 2019 fue igual a 3.248 millones de euros, mientras que el montante total de transferencias emitidas fue de 5.407 millones de euros. En el año 2020, estas cifras fueron de 3.163 millones de euros y 5.554 millones de euros, respectivamente. Por tanto, observamos que en España los hogares emiten más transferencias a otros hogares de las que reciben. Esto va a generar que la renta privada sea, en general, más reducida que la renta de mercado.

Cabe mencionar que existen otros análisis sobre la distribución de la renta que incluyen otros conceptos, como, por ejemplo, el caso de las transferencias públicas en especie como la sanidad (Álvarez, 2019) o la educación (Bautista, 2020), así como la utilización de otras variables como pudiera ser el alquiler imputado. No obstante, una imputación territorializada de esas transferencias en especie sanitarias o educativas escapa de las dimensiones de este estudio. Por consiguiente, en este trabajo nos centraremos más en las diferencias de desigualdad entre territorios que en la precisión territorial de la medida de desigualdad de cada uno de ellos.

En este punto ya se ha expuesto cómo se obtienen los valores de las distintas variables que se van a utilizar a lo largo del trabajo a nivel de hogar. Ahora bien, el análisis que se va a llevar a cabo posteriormente se podría realizar tomando los hogares como unidad de medida o considerando los individuos. En trabajos anteriores, se ha observado que el nivel de desigualdad no es especialmente diferente cuando se analiza tomando el hogar como unidad de medida (Anghel et al., 2018); no obstante, en este trabajo, siguiendo la línea de Bautista (2020), Araoz (2024) o la forma de presentar los resultados del INE, se ha optado por emplear los individuos como elemento de estudio por considerarse más adecuado. Sin embargo, esto plantea ciertas cuestiones acerca de cómo trasladar a cada individuo la renta a nivel hogar calculada previamente.

La primera opción que podría aplicarse sería repartir la renta del hogar de forma igualitaria entre todos los individuos de este. Esta asunción, tal y como señala Goerlich (2016), implicaría asumir que un hogar donde viven seis personas necesita tres veces más renta para alcanzar el mismo nivel de vida que un hogar donde solo residan dos personas.

Asimismo, tal y como se muestra en la siguiente ilustración, la renta de los hogares aumenta con respecto a su tamaño debido probablemente a que más miembros pueden aportar mayores ingresos o a que determinadas ayudas se establecen en función del número de individuos que residen en el hogar, si bien, este crecimiento se va reduciendo. Por ello, la renta per cápita se reduce conforme aumenta el número de miembros que reside en el hogar.



Fuente: Elaboración propia.

Este hecho hace que la renta per cápita sea un indicador discutible del nivel de vida de los hogares cuando estos presentan diferente número de miembros. Para solventar este problema lo habitual es introducir ponderaciones para ajustar la renta de los hogares en función del número de individuos y sus características. A estas transformaciones se las conoce como escalas de equivalencia, si bien no existe un consenso acerca de qué ajuste realizar (Goerlich y Villar, 2009). Cabe destacar a su vez que la literatura evidencia que la elección de la escala tiene un impacto decisivo en los resultados obtenidos (Buhmann et al., 1988 y Jenkins y Lambert, 1993).

En definitiva, la cuestión es definir una función de nivel de vida del individuo que tenga en cuenta tanto la renta de su hogar como el número de individuos que lo componen de la siguiente forma:

$$y_{i,h} = m(y_h, n_h) \quad (9)$$

Donde $y_{i,h}$ es la renta del individuo i que pertenece al hogar h y $m(\cdot)$ es una función que depende de la renta del hogar y_h y del número de individuos del mismo n_h .

Ahora bien, a la hora de determinar la función, esta ha de cumplir una serie de características deseables tal y como señala Bautista (2020). En primer lugar, ha de ser no homogénea de grado cero (ni inferior) en individuos y renta de los hogares. Esto garantiza que la renta de los individuos depende del número de individuos y de la renta y no solo de sus proporciones, al mismo tiempo que la renta de los individuos no se reduce cuando estas dos variables aumentan proporcionalmente. En definitiva, ha de cumplir lo siguiente:

$$m(\lambda y_h, \lambda n_h) > m(y_h, n_h) \quad \forall \lambda > 1 \quad (10)$$

Esta condición es importante desde un punto de vista económico ya que estamos asumiendo que existen una serie de costes fijos en los hogares por lo que, si aumentan en la misma proporción los individuos y la renta, cada individuo podrá disfrutar de más renta al poder repartir los costes.

Por otra parte, el nivel de vida de un individuo tiene que crecer cuando aumenta la renta del hogar y el número de individuos se mantiene constante. De la misma manera, si la renta del hogar se mantiene pero el número de individuos se incrementa, la renta de la que dispondrá cada individuo de manera individual se verá reducida. Por tanto, hemos de imponer las siguientes restricciones:

$$\begin{aligned}
i) \quad & \frac{\partial m(y_h, n_h)}{\partial y_h} > 0 \\
ii) \quad & \frac{\partial m(y_h, n_h)}{\partial n_h} < 0
\end{aligned} \tag{11}$$

Una vez se han establecido las condiciones básicas que deben tener las escalas de equivalencia, a continuación se muestran algunas de las formulaciones planteadas por la literatura.

Así, existe una corriente (Buhmann et al., 1988) que ha optado por utilizar escalas donde la premisa principal es la elasticidad de las necesidades en función del tamaño del hogar, entendiendo esta como constante e independiente respecto del resto de características del hogar (Atkinson, Rainwater y Smeeding, 1995). Así podríamos definir, por ejemplo:

$$m(y_h, n_h) = \frac{y_h}{n_h^\theta} \quad \theta \in [0,1] \tag{12}$$

En este caso, θ corresponde con la elasticidad de las necesidades respecto de cada individuo del hogar. En el caso de que la elasticidad fuera igual a uno, estaríamos representando la renta per cápita y no se tendrían en cuenta las economías de escala que podemos asociar al tamaño del hogar. En el caso de la OCDE, se utiliza una elasticidad igual a 0.5, por lo que se considera que un hogar compuesto por una única persona tiene la mitad de necesidades que un hogar compuesto por cuatro individuos.

Por otra parte, organismos como Eurostat (Balestra y Oehler, 2023) o el INE (2023) y trabajos como Bautista (2020) y Goerlich (2016) establecen que no cabe considerar que los individuos adultos tienen las mismas características que los niños. Por tanto, se establece la siguiente escala:

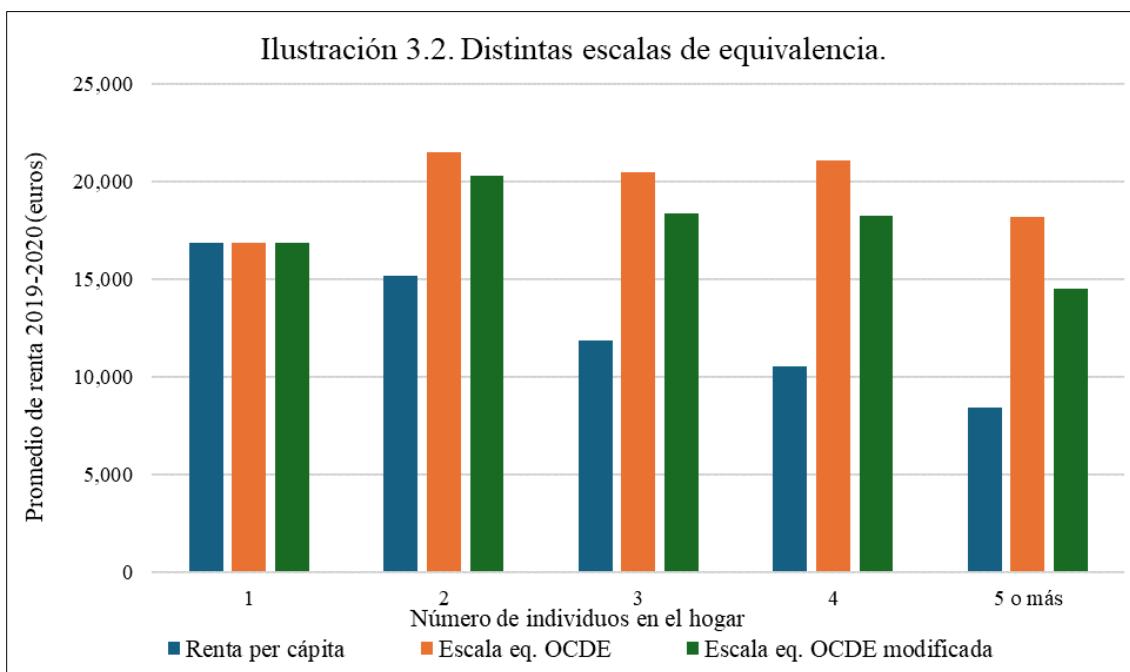
$$m(y_h, n_a, n_n) = \frac{y_h}{1 + \alpha(n_a - 1) + \beta n_n} \tag{13}$$

Donde n_a representa el número de adultos del hogar menos uno y n_n el número de niños. En particular, en los trabajos mencionados anteriormente se adopta lo que se conoce como escala de la OCDE modificada (Hagenaars, De Vos y Zaidi, 1994). En esta escala se considera que el primer miembro de la familia tiene valor igual a uno, cada adulto adicional tiene un valor de 0.5 y cada niño tiene un valor de 0.3. Entendiendo

adultos como aquellos individuos con 14 o más años. Por tanto, se le está asignando un valor a α igual a 0.5 y a β igual a 0.3.

De esta formulación, tal y como señala Goerlich (2016), cabe destacar que el ajuste realizado es en cierta medida arbitrario y que se puede estudiar cómo varían los resultados en función de las escalas con solo variar los valores de los parámetros.

Al igual que en la ilustración 3.1 se apreciaba la diferencia entre renta del hogar y renta per cápita, a continuación se muestra la diferencia existente entre la renta per cápita y las distintas rentas equivalentes según las escalas de equivalencia anteriormente presentadas:



Fuente: Elaboración propia.

De acuerdo con la ilustración anterior, podemos apreciar la escala con mayor estabilidad es la de la OCDE modificada utilizada por Eurostat, si bien no existe una diferencia excesiva entre esta y la de la OCDE.

En este trabajo se ha decidido utilizar la escala de la OCDE modificada por cumplir con los requisitos deseables y permitir la comparabilidad con otros trabajos similares.

En definitiva, en este trabajo se van a obtener las distintas rentas de cada individuo aplicando la siguiente expresión:

$$y_{i,h} = \frac{y_h}{1 + 0.5(n_a - 1) + 0.3n_n} = \frac{y_h}{HX240} \quad (14)$$

Siendo, tal y como se ha mencionado previamente, $y_{i,h}$ la renta equivalente considerada del individuo i que pertenece al hogar h , y_h la renta considerada del hogar h , n_a el número de individuos con 14 o más años del hogar h y n_n el número de individuos menores de 14 años del hogar h . Cabe destacar que, tal y como se muestra en la ecuación 14, el denominador es igual a la variable HX240 presente en los ficheros de metadatos de la ECV (INE, 2021a), habiéndose comprobado que el cálculo se realiza de forma adecuada.

Finalmente, cabe destacar que en los ficheros de la ECV proporcionados por el INE, la muestra está compuesta por 38.011 individuos y 15.043 hogares en el año 2019 y 52.325 individuos y 21.007 hogares en el año 2020. Con la finalidad de obtener resultados poblacionales en lugar de muestrales, es decir, llegar a 46.874.430 individuos y 18.773.903 hogares en el año 2019 y 46.877.576 individuos y 18.825.525 hogares en el año 2020, resulta necesario utilizar los factores de elevación proporcionados en la propia encuestar, esto es emplear las variables RB050 para individuos y DB090 para hogares (INE, 2021a). Este tratamiento, que no se realiza en los trabajos de Bautista (2020) o Araoz (2024), sí que es utilizado por organismos como Eurostat a la hora de calcular sus índices (Eurostat, 2023).

4-. Resultados

Una vez se han definido tanto los indicadores a utilizar como la metodología empleada para obtener las distintas variables, el objetivo de este apartado es tratar de analizar los resultados obtenidos en España. Cabe destacar que los cálculos se han realizado a través del software libre R (R Core Team, 2023), adjuntándose el código utilizado en el Anexo I para garantizar la replicabilidad de los resultados.

4.1-. Efectos del sector público en las distintas regiones españolas

En primer lugar, resulta de interés poner el foco en cómo funciona la labor redistributiva del sector público en España, atendiendo en particular a las posibles diferencias entre las distintas comunidades autónomas.

Para ello, se ha optado por tomar los datos del año 2019 (INE, 2021b) dado que son los últimos disponibles a fecha de realización de este trabajo no influenciados por la pandemia COVID-19, por lo que nos permiten obtener una visión de la situación habitual española.

En primer lugar, se han calculado los índices de Gini para cada una de las cuatro variables de renta definidas en el capítulo anterior para cada una de las 17 comunidades autónomas y 2 ciudades autónomas, así como para el conjunto de España. Los resultados obtenidos se muestran a continuación:

Tabla 4.1. Índices de Gini para distintas variables de renta por Comunidades Autónomas (año 2019)

Comunidad Autónoma	Renta de Mercado	Renta Privada	Renta Bruta	Renta Disponible
Andalucía	50.05	50.18	35.34	30.78
Aragón	44.81	44.94	31.23	27.44
Canarias	50.23	50.49	37.53	33.10
Cantabria	51.64	51.67	35.58	30.81
Castilla y León	47.60	47.66	32.62	27.86
Castilla-La Mancha	46.51	46.48	33.25	28.95
Cataluña	48.50	48.44	35.99	31.65
Ciudad Autónoma de Ceuta	54.35	54.43	42.45	40.03
Ciudad Autónoma de Melilla	57.20	56.80	46.36	43.90
Comunidad de Madrid	47.15	47.15	36.75	32.91
Comunidad Foral de Navarra	44.33	44.46	31.19	27.28
Comunidad Valenciana	49.15	49.26	36.23	32.30
Extremadura	50.38	50.86	32.75	28.00
Galicia	49.68	49.71	33.88	29.71
Illes Balears	38.41	38.44	29.97	26.46
La Rioja	44.22	43.92	30.97	28.21
País Vasco	47.31	47.15	31.93	28.77
Principado de Asturias	56.69	56.52	37.71	34.00
Región de Murcia	45.95	46.04	32.60	28.82
España	49.42	49.47	36.34	32.09

Tal y como se desprende de la tabla anterior, los índices de Gini obtenidos tanto para la renta de mercado como para la renta privada son notablemente superiores a los obtenidos en la renta bruta y la renta disponible, lo que nos evidencia como el sector público contribuye a reducir las desigualdades entre los individuos en España y, en particular, en cada una de las comunidades autónomas.

En concreto, los índices de las dos primeras variables se sitúan mayoritariamente en valores por encima de 40 y en siete de ellas por encima de 50. Cabe destacar, tal y como se podía anticipar, que los resultados obtenidos para estas dos mediciones son muy similares pues la única diferencia entre ellas son las transferencias realizadas entre los hogares y este importe es relativamente pequeño en comparación con el resto de rentas.

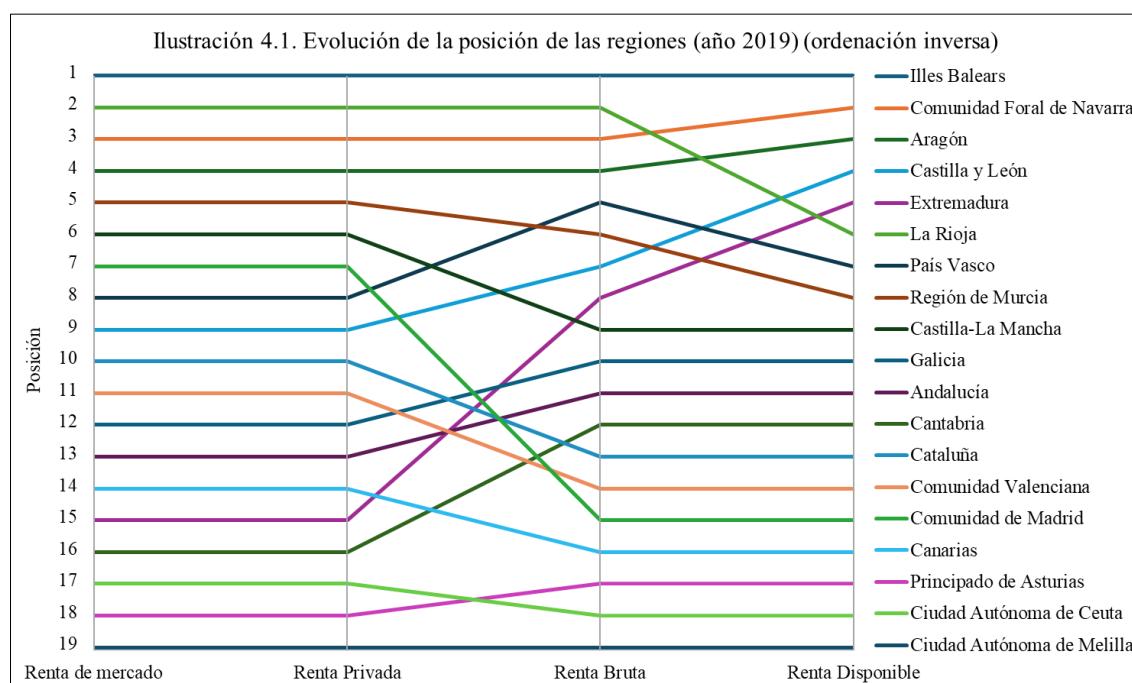
En el caso de la renta bruta, los índices de Gini obtenidos toman valores entre 30 y 40, observando notables descensos al introducir las transferencias monetarias. Para la renta disponible, los valores se sitúan mayoritariamente entre 25 y 35, siendo el descenso, comparando con la renta bruta, menor que el salto anterior.

Por otra parte, atendiendo a medidas de dispersión entre las distintas comunidades autónomas, se ha obtenido que la variable con más varianza es la renta de mercado, seguida de la renta privada, la renta disponible y, finalmente, la renta bruta. Por tanto,

podemos concluir que las transferencias monetarias del sector público disminuyen las diferencias de desigualdad existentes entre regiones más que los impuestos directos, aunque estos también contribuyen a esta labor.

El siguiente análisis que se va a realizar busca estudiar la posición en la que se encuentra cada Comunidad Autónoma en cada una de las variables de renta que se consideran. En este sentido, lo esperable sería que se respetara el principio de ordinalidad, es decir, que si una región es la menos desigual atendiendo a su renta privada, siguiera siendo la menos desigual cuando analizamos su renta disponible.

A continuación, se muestra un gráfico donde se ordenan de manera inversa las regiones en función de la posición que ocupan en relación con el grado de desigualdad según cada variable con la finalidad de estudiar su evolución según la renta que se considere:



Fuente: Elaboración propia.

De los resultados anteriores la primera cuestión que cabe resaltar es que, al igual que se ha expuesto anteriormente, el orden de las comunidades autónomas atendiendo a la renta de mercado es igual al orden obtenido de acuerdo con la renta privada. En estas variables, las tres regiones con más desigualdad son la Ciudad Autónoma de Melilla, el Principado de Asturias y la Ciudad Autónoma de Ceuta. En el otro extremo de la lista se sitúan Illes Balears, La Rioja, la Comunidad Foral de Navarra y Aragón.

Ahora bien, tal y como se observa en el gráfico, los principales cambios en orden se dan cuando se incluyen las transferencias en los ingresos de los hogares, es decir, en el paso de renta de mercado a renta bruta. Se observa que todas las regiones cambian de orden con excepción de las cuatro menos desiguales (Illes Balears, La Rioja, Comunidad Foral de Navarra y Aragón) y la más desigual (Ciudad Autónoma de Melilla).

Respecto de las comunidades con más cambios, destacan en cuanto a reducir la desigualdad Extremadura y Cantabria, subiendo 7 y 4 puestos, respectivamente. Por otra parte, en el ámbito opuesto (descender puestos en el ranquin de desigualdad) sobresale la Comunidad de Madrid que baja 8 puestos.

Finalmente, cuando se incluyen los impuestos directos, en el tránsito de renta bruta a renta disponible, no se observa ningún cambio ni en la región menos desigual que sigue siendo Illes Balears, ni en las 11 comunidades más desiguales. Destaca el caso de La Rioja que cae 4 posiciones y tanto Extremadura como Castilla y León subiendo suben 3 posiciones cada una.

Analizando las diferencias de forma agregada, es decir, desde la renta que obtienen los individuos fruto de su relación con el mercado ya sea a través de su fuerza de trabajo o de sus ganancias de capital frente a la renta de la que disponen finalmente una vez se han tenido en cuenta tanto las transferencias monetarias públicas que reciben como los impuestos directos pagados, únicamente dos comunidades se mantienen en la posición de origen. Resulta reseñable que estas comunidades son la más desigual, la Ciudad Autónoma de Melilla y la menos desigual, Illes Balears. En cuanto a cambios, mejora notablemente su posición Extremadura subiendo 10 puestos y empeora fuertemente la Comunidad de Madrid bajando 8.

Las razones de estas variaciones son diversas y multifactoriales. Por un lado, hay que tener en cuenta que el principal factor reductor de la desigualdad en España son las transferencias monetarias públicas y, en particular, las pensiones (López Laborda et al., 2023). Por lo tanto, no es de extrañar que aquellas regiones con mayor porcentaje de población jubilada vean un efecto del sector público en la reducción de la desigualdad mayor. De la misma manera, regiones como Madrid, con un mayor dinamismo económico podrían no estar utilizando tanto este tipo de mecanismos y, aunque ven reducida su desigualdad por la acción del sector público, lo hacen en menor medida.

Ahora bien, tras este análisis cualitativo basado en la posición que ocupan las distintas regiones, se va a realizar un estudio de carácter más cuantitativo mediante el uso del índice de Reynolds-Smolensky definido en epígrafes anteriores.

A continuación, se muestran tanto los índices de Gini y concentración como los índices de Reynolds- Smolensky obtenidos para cada una de las regiones en el año 2019:

Tabla 4.2. Índices de Reynolds-Smolensky para distintas variables de renta por Comunidades Autónomas (año 2019)

Comunidad Autónoma	Índice de Gini/Concentración			Índice de Reynolds-Smolensky		
	Renta Privada (1)	Renta Privada + Transf. (2)	Renta Disponible (3)	Transferencias monetarias (1-2)	Impuestos directos (2-3)	Total (1-3)
Andalucía	50.18	29.65	21.33	20.53	8.32	28.84
Aragón	44.94	26.50	18.46	18.44	8.04	26.47
Canarias	50.49	33.99	26.70	16.50	7.28	23.79
Cantabria	51.67	29.76	21.42	21.90	8.35	30.25
Castilla y León	47.66	27.48	19.27	20.19	8.21	28.39
Castilla-La Mancha	46.48	29.50	22.09	16.98	7.40	24.38
Cataluña	48.44	31.75	23.73	16.69	8.02	24.71
Ciudad Autónoma de Ceuta	54.43	37.05	32.04	17.38	5.01	22.39
Ciudad Autónoma de Melilla	56.80	43.64	39.82	13.15	3.82	16.98
Comunidad de Madrid	47.15	32.59	25.44	14.56	7.15	21.71
Comunidad Foral de Navarra	44.46	27.09	19.69	17.37	7.40	24.77
Comunidad Valenciana	49.26	30.73	22.98	18.53	7.75	26.28
Extremadura	50.86	27.51	19.19	23.34	8.32	31.66
Galicia	49.71	29.02	20.96	20.69	8.06	28.76
Illes Balears	38.44	27.18	21.01	11.26	6.18	17.43
La Rioja	43.92	26.66	20.03	17.26	6.63	23.89
País Vasco	47.15	26.02	19.05	21.13	6.97	28.10
Principado de Asturias	56.52	30.89	22.91	25.63	7.98	33.61
Región de Murcia	46.04	27.51	19.36	18.53	8.15	26.68
España	49.47	31.38	23.51	18.09	7.87	25.96

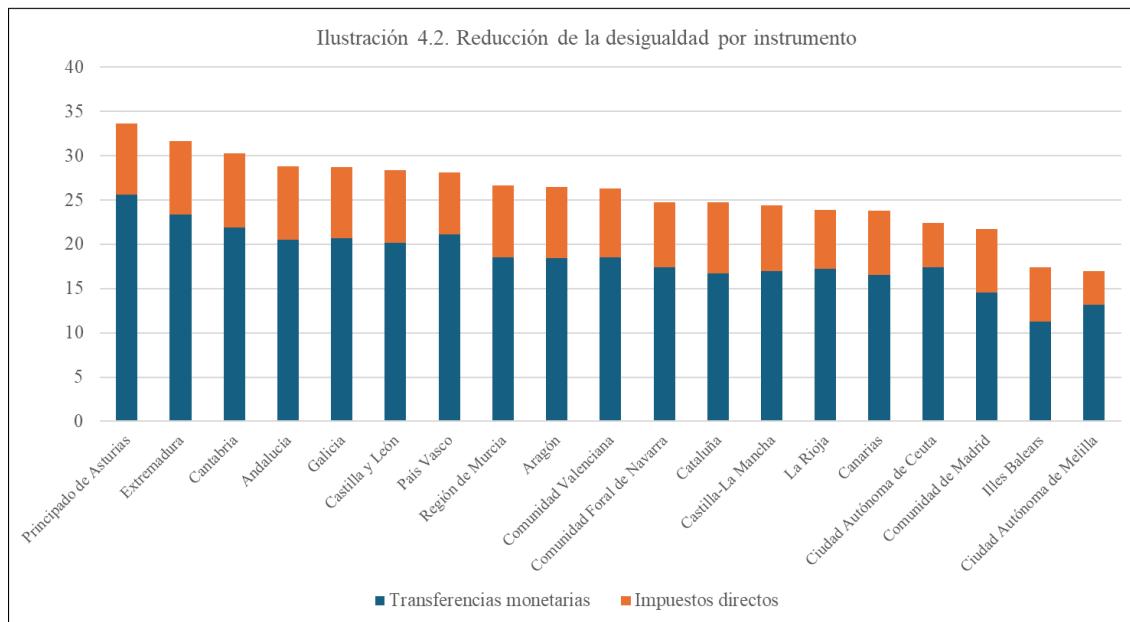
Fuente: Elaboración propia a partir de la ECV.

La primera conclusión que se puede extraer de la tabla anterior es la labor redistributiva del estado se lleva a cabo principalmente a través de las transferencias monetarias públicas, representando en torno al 70% del total. Asimismo, este resultado se produce en todas las comunidades autónomas, en particular, la región donde tiene un mayor efecto redistributivo es el Principado de Asturias, seguido de Extremadura y Cantabria. En el lado contrario, las regiones donde las transferencias monetarias tienen un menor impacto redistributivo son la Ciudad Autónoma de Melilla, Illes Balears y la Comunidad de Madrid.

En relación con el efecto redistributivo de los impuestos directos, las comunidades autónomas donde este efecto es más fuerte son Cantabria, Extremadura y Andalucía. Por el contrario, las regiones donde menor es el impacto son la Ciudad Autónoma de Melilla, la Ciudad Autónoma de Ceuta e Illes Balears.

Hay que tener en cuenta que cuanto mayor sea la desigualdad de partida de la renta privada mayor será el efecto redistributivo y viceversa. Por ello, no es de extrañar que en Illes Balears el efecto redistributivo sea tan bajo en comparación con otras regiones.

Con la finalidad de estudiar en profundidad esta cuestión, poniendo especial énfasis en las comunidades autónomas con mayor efecto redistributivo por parte del sector público, se muestra el siguiente gráfico donde se representa la reducción de la desigualdad en cada una de las comunidades ordenadas de mayor a menor efecto redistributivo.



Fuente: Elaboración propia.

Atendiendo al gráfico anterior, se observa que las comunidades donde el sector público tiene un mayor efecto redistributivo son el Principado de Asturias, Extremadura y Cantabria. Por el contrario, aquellas comunidades donde la acción del estado tiene un menor impacto en la reducción de la desigualdad son la Ciudad Autónoma de Melilla, Illes Balears y la Comunidad de Madrid. Es importante señalar de nuevo que este impacto se puede ver afectado por el punto de partida, en este caso, la desigualdad en la renta privada, no obstante, se aprecia un patrón claro en el sentido de que aquellas comunidades con mayor porcentaje de pensionistas son aquellas donde el sector público tiene mayor capacidad redistributiva. Esto se debe a que, como se ha mencionado previamente, buena parte del poder redistributivo del sector público español viene de las pensiones (López Laborda et al., 2023).

Analizando ahora en qué regiones tienen más peso las transferencias monetarias públicas en términos relativos estas serían las ciudades autónomas de Ceuta y Melilla con un peso en la redistribución del 77%. En el lado opuesto, aquellas comunidades donde más impacto tienen los impuestos directos son Illes Balears y la Comunidad de Madrid con pesos del 35% y 33%, respectivamente.

Esto nos está indicando que aquellas regiones con un mayor porcentaje de población activa y menor de pensionistas, las que podríamos considerar más dinámicas económicamente, tienen un mayor porcentaje de redistribución vía impuestos directos respecto al resto de comunidades, pese a que la redistribución vía transferencias monetarias sigue siendo la principal fuente redistributiva en España.

En definitiva, se puede afirmar que en España la reducción de la desigualdad se da fundamentalmente por la vía de las transferencias monetarias públicas. Estos resultados son similares a los obtenidos en otros trabajos para el caso español (Bautista, 2020 y López Laborda et al., 2023), así como para el resto de economías desarrolladas (FMI, 2014 y Joumard et al., 2013). No obstante, el hecho de que la redistribución vía transferencias monetarias se produzca fundamentalmente a través de las pensiones públicas es una característica idiosincrática del caso español.

4.2-. Efectos del sector público en la pandemia Covid-19

Una vez se ha analizado el efecto que tiene sobre la desigualdad la acción sector público en España analizando, en particular, como contribuye en esta tarea en cada una de las comunidades autónomas, en este apartado se busca analizar qué impacto ha tenido en un contexto particular, la pandemia COVID 19.

Con la finalidad de gestionar la situación originada, el Gobierno de España decretó el Estado de Alarma el 14 de marzo de 2020, que finalmente se alargó hasta el 21 de junio. Este decreto supuso la implantación de diversas medidas restrictivas sin precedentes como la limitación de la movilidad y el confinamiento en el domicilio para toda la población, permitiéndose tan solo las salidas para ir a trabajar (en el caso de los trabajos considerados esenciales), comprar artículos de primera necesidad, acudir a citas médicas, volver a la residencia habitual y asistir a personas dependientes, aparte de otras causas de fuerza mayor (RDL 463/2020, de 14 de marzo).

Aunque estas medidas consiguieron reducir de forma efectiva la incidencia de la COVID-19 en el país, generaron una notable caída en los principales indicadores económicos. Según datos del INE, el PIB cayó un 11% en el año 2020, la tasa de paro subió hasta el 16.3% en el tercer trimestre y el porcentaje de trabajadores en Expedientes de Regulación Temporal de Empleo (ERTE) llegó a superar el 20%.

Cabe destacar que este análisis se va a llevar a cabo de forma agregada para el conjunto de España con la finalidad de eliminar posibles distorsiones regionales. En primer lugar, se ha realizado un estudio para tratar de ver cuál es el promedio de cada una de las rentas calculadas por unidad de consumo, tanto en 2019 como en 2020 por tramos de renta bruta, con la finalidad de detectar qué cambios se han producido en los ingresos de los individuos como consecuencia de la pandemia. Los resultados obtenidos se muestran a continuación:

Tabla 4.3. Promedio de distintas variables de renta por unidad de consumo en España (año 2019)

Tramos	Renta Bruta	Renta de Mercado	Renta Privada	Renta Bruta	Renta Disponible	Saldo
Quintil 1	3,844.88	3,891.16	7,730.84	6,670.01	2,778.85	
Quintil 2	10,231.88	10,225.39	15,581.26	12,421.30	2,195.91	
Quintil 3	16,500.13	16,403.55	22,481.79	16,639.71	236.16	
Quintil 4	26,056.77	25,993.55	32,052.09	22,165.97	-3,827.58	
Decil 9	37,287.74	37,164.69	43,515.48	28,646.68	-8,518.01	
Centil 91-99	56,185.56	55,982.50	62,101.05	38,643.29	-17,339.21	
Centil 100	118,646.20	117,974.82	126,332.01	77,457.54	-40,517.28	
Total	20,383.94	20,328.27	25,809.01	18,115.61	-2,212.66	

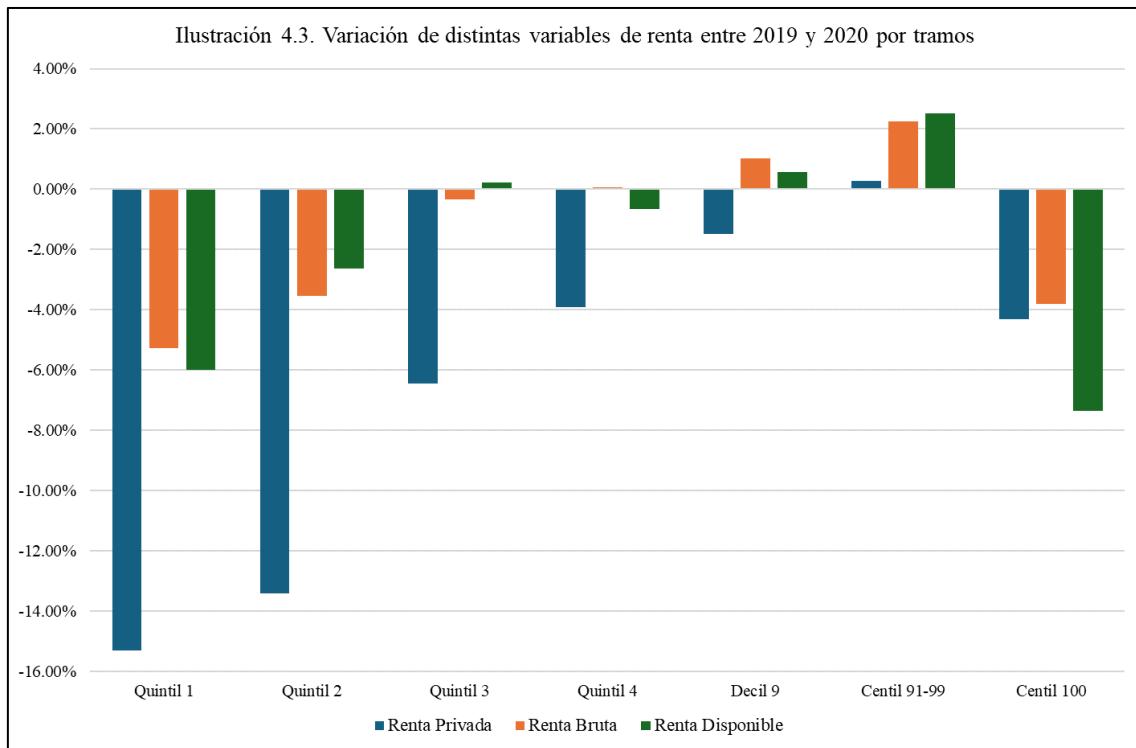
Fuente: Elaboración propia a partir de la ECV.

Tabla 4.4. Promedio de distintas variables de renta por unidad de consumo en España (año 2020)

Tramos	Renta Bruta	Renta de Mercado	Renta Privada	Renta Bruta	Renta Disponible	Saldo
Quintil 1	3,222.55	3,295.34	7,322.20	6,270.27	2,974.93	
Quintil 2	8,843.25	8,852.74	15,030.92	12,094.91	3,242.17	
Quintil 3	15,393.20	15,347.50	22,406.08	16,675.64	1,328.14	
Quintil 4	25,084.85	24,974.87	32,071.49	22,019.92	-2,954.95	
Decil 9	36,779.81	36,612.02	43,958.31	28,806.76	-7,805.26	
Centil 91-99	56,429.97	56,135.71	63,505.70	39,616.63	-16,519.08	
Centil 100	113,876.10	112,892.84	121,524.70	71,751.13	-41,141.71	
Total	19,670.67	19,608.61	25,905.02	18,103.00	-1,505.61	

Fuente: Elaboración propia a partir de la ECV.

Asimismo, se han calculado las variaciones relativas de la renta privada, la renta bruta y la renta disponible de los distintos tramos de renta representados en las tablas anteriores entre los años 2020 y 2019. Los resultados se pueden observar en la siguiente ilustración:



Fuente: Elaboración propia.

El primer resultado que se desprende de la ilustración es que, atendiendo al conjunto de los individuos, la renta privada ha disminuido un 3.5%, la renta bruta ha aumentado un 0.37% y la renta disponible ha disminuido un 0.07%. Por tanto, se concluye que la intervención pública ha provocado que la renta disponible permanezca prácticamente estable frente al descenso que cabría esperar por el efecto de la pandemia.

Analizando cómo han evolucionado las rentas privada y de mercado por tramos de renta, las principales caídas se dan en los dos primeros quintiles con un descenso superior al 10%. Esta bajada da en todos los tramos con las excepción de los individuos situados entre los percentiles 91 y 99 que han visto su renta privada aumentar un 0.27% en el año, lo cual no deja de resultar sorprendente.

En cuanto a la renta bruta, se siguen observando caídas en los tres primeros quintiles así como en el percentil 100. En el resto de tramos se aprecia un crecimiento de la renta bruta, destacando otra vez los percentiles 91-99 con un crecimiento del 2.26%.

Cuando se atiende a la renta disponible, los únicos tramos con crecimiento son el decil 9 y los percentiles 91-99 y, aunque se mantiene prácticamente estable, el quintil 3. Por tanto, observamos como estos segmentos de población han visto incrementada su renta disponible a pesar de la crisis pandémica. En el lado opuesto, los grupos de

población donde se observa un mayor descenso de la renta son el primer quintil con una caída del 5.99% y el 1% con más renta que ha sufrido una caída del 7.37%.

Finalmente, la variable saldo, nos indica la diferencia entre la renta disponible y la renta privada. Por tanto nos está dando información acerca del montante neto que los individuos han obtenido/contribuido como consecuencia de la intervención pública. El primer resultado es que de media los individuos pagaban 2,212.66€ al sector público en 2019 frente a 1,505.62€ en el año 2020. Por tanto, la contribución media de los individuos se ha reducido un 31.95%.

Adicionalmente, en ambos años, más del 60% de la población ha recibido más transferencias monetarias de lo que ha contribuido en impuestos directos al sector público existiendo además monotonía en el sentido de que conforme aumenta la renta más negativo es el saldo. Ahora bien, mientras que el primer quintil ha aumentado su saldo en un 7.06%, el segundo quintil lo ha hecho en un 47.65% y el tercer quintil lo ha multiplicado por más de cinco. El cuarto quintil también ha sido notablemente beneficiado reduciendo su saldo negativo un 22.8%, el resto de tramos han sido beneficiados comparando 2019 con 2020, aunque en menor medida. Finalmente, los individuos pertenecientes al 1% con más renta han visto su saldo negativo incrementarse en un 1.54%.

En definitiva, este análisis ex post permite concluir que durante la pandemia los principales beneficiados han sido las rentas situadas en entre los percentiles 91 y 99 de la distribución.

Una vez realizado el análisis anterior, se han calculado los valores del índice de Gini para las distintas variables de renta tanto para el año 2019 como para el año 2020, obteniendo los siguientes resultados:

Tabla 4.5. Índices de Gini para distintas variables de renta en España

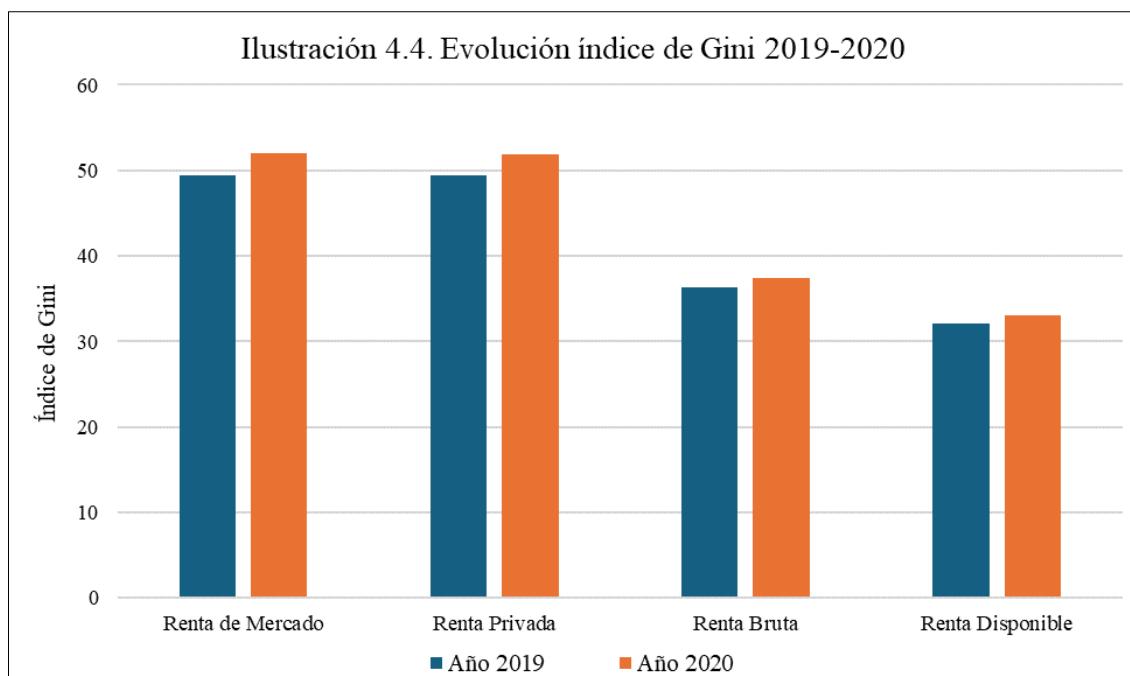
Año	Renta de Mercado	Renta Privada	Renta Bruta	Renta Disponible
2019	49.42	49.47	36.34	32.09
2020	51.95	51.94	37.37	33.02

Fuente: Elaboración propia a partir de la ECV.

La primera conclusión que se puede extraer de la tabla anterior es que tanto en la renta de mercado como en la renta privada se observa un incremento de la desigualdad

de en torno a un 5% producida por la situación pandémica. Ahora bien, si miramos la renta bruta, el índice de Gini se incrementa únicamente un 2.8% y, finalmente, en el caso de la renta disponible, el incremento es del 2.9%. Por tanto, estamos viendo que no solo el estado contribuye a reducir la desigualdad, sino que el diseño anticíclico de ciertas transferencias públicas como el subsidio por desempleo o las prestaciones de supervivencia así como las medidas concretas establecidas durante la pandemia como el ingreso mínimo vital y los ERTE hicieron que la desigualdad de la renta disponible aumentara menos de lo que aumentó la desigualdad de mercado.

A continuación se muestra una ilustración donde se puede visualizar de forma paralela la evolución de las magnitudes:



Fuente: Elaboración propia.

La principal conclusión que se puede obtener de la ilustración anterior es que el salto existente entre la renta privada y la renta bruta es mayor que el visto entre la renta bruta y la renta disponible tanto en el año 2019 como en el año 2020, por lo que se mantiene la tendencia vista en el epígrafe anterior.

Cabe destacar que la literatura sobre la evolución de los índices de Gini durante la pandemia presenta resultados mixtos, con ciertos trabajos en línea con lo calculado en el presente documento (aumento de la desigualdad) como podría ser Bandrés Moliné

(2023) y, por el otro lado, trabajos como López Laborda et al. (2023) que obtienen una reducción de la desigualdad entre 2019 y 2020.

Estas diferencias en los cálculos de los índices de Gini entre distintos trabajos se han venido dando en la literatura debido a las múltiples decisiones metodológicas que resulta necesario abordar. No obstante, en este caso, las divergencias resultan más relevantes porque la diferencia en los cálculos nos puede llevar a conclusiones contrapuestas, aumento o reducción de la desigualdad. Por tanto, puede resultar relevante para posteriores investigaciones tratar de unificar las metodologías y establecer el criterio más adecuado.

Con la finalidad de profundizar en cuales son los instrumentos del sector público que más contribuyen a la reducción de la desigualdad se procede a calcular, al igual que en el epígrafe anterior, el índice de Reynolds-Smolensky, tal como muestra la tabla siguiente:

Tabla 4.6. Índices de Reynolds-Smolensky para distintas variables de renta en España

Año	Transferencias públicas	Impuestos directos	Total
2019	18.09	7.87	25.96
2020	19.92	7.91	27.83

Fuente: Elaboración propia a partir de la ECV.

Tal y como se desprende de la tabla anterior, ha aumentado el poder redistributivo tanto de las transferencias monetarias públicas como de los impuestos directos. Lo que resulta reseñable es que el poder redistributivo de las transferencias públicas ha aumentado un 10.13% mientras que el de los impuestos directos únicamente un 0.49%. Este resultado es esperable ya que, como se ha mencionado previamente, las medidas que se pusieron en marcha durante la pandemia venían más por el lado de las transferencias monetarias que por el de los impuestos directos.

Asimismo, se aprecia como, si en 2019 las transferencias representaban el 69.69% del poder redistributivo del sector público español, en el año 2020 esta cifra ascendió hasta el 71.52%.

En resumen, tras esta evaluación *ex post* se puede concluir que entre 2019 y 2020 ha habido un aumento de la capacidad redistributiva del sector público. Este resultado está en línea con otros trabajos publicados como López Laborda et al. (2023).

5-. Conclusiones

En este trabajo se ha estudiado el impacto de la acción del sector público en la desigualdad poniendo énfasis en distinguir qué herramientas contribuyen más a esa función redistributiva al tiempo que se ha testeado si es homogénea a lo largo del territorio y qué impacto tuvo la pandemia en el esfuerzo redistributivo.

Para cumplir estos objetivos que nos marcábamos en la introducción, se han tomado datos públicos procedentes de la Encuesta de Condiciones de Vida elaborada anualmente por el INE. Posteriormente, se han analizado diferentes metodologías propuestas en la literatura para obtener la renta poniendo énfasis en sus diferencias hasta obtener una metodología que permite imputar de manera individual la renta a los individuos de forma adecuada.

En relación con la capacidad redistributiva del sector público, se ha evidenciado, tal y como venía afirmando la literatura, que es mucho más fuerte por parte de las transferencias monetarias que por parte de los impuestos directos. Dicha característica se mantiene cuando el análisis se traslada a las distintas comunidades autónomas, si bien, la capacidad redistributiva es distinta entre ellas y se observa que la ordenación de las comunidades en función de la desigualdad varía tras la acción pública. Se detecta, además, que en las comunidades más envejecidas es donde mayor peso tienen las transferencias debido al fuerte impacto redistributivo de las pensiones, en línea con lo reflejado en la literatura.

Para el estudio del efecto redistributivo en la pandemia, hemos podido realizar un análisis ex post de la acción del sector público. En primer lugar, hemos observado que la mayoría de los individuos vieron sus rentas descender entre 2019 y 2020 como consecuencia de la pandemia con excepción de los individuos pertenecientes al quintil 3, al decil 9 y los situados entre los percentiles 91 y 99 de la distribución, siendo estos últimos los más beneficiados. Cuando se ha analizado la evolución de la desigualdad, esta aumentó, pero se demuestra que lo hizo menos debido a la acción del sector público. Además, se observa como el poder redistributivo del Estado aumentó en dos puntos porcentuales.

Adicionalmente, una ventaja fundamental del trabajo es su replicabilidad debido a la utilización de un software libre como es R en todo el proceso y la presentación de los códigos utilizados en el Anexo I.

En este sentido, el presente estudio tiene un recorrido de ampliación de diversas formas. En primer lugar, el análisis efectuado a nivel regional se podría ampliar a un número superior de períodos con la finalidad de estudiar si las conclusiones obtenidas son tendencias persistentes como parece o detectar posibles cambios estructurales en algunas regiones. En segundo lugar, dado que la metodología empleada parte de una encuesta que forma parte de EU-SILC, se podría trasladar el estudio al resto de países europeos que siguen este estándar.

Asimismo, futuras investigaciones podrían plantear la utilización de diferentes índices, además de los trabajados en este documento, para verificar si las conclusiones obtenidas se mantienen, teniendo en cuenta los problemas previamente mencionados en el índice de Gini.

6-. Referencias bibliográficas

- Acemoğlu, D., Naidu, S., Restrepo, P. y Robinson, J.A. (2015). Democracy, Redistribution, and Inequality. En Atkinson, A.B. y Bourguignon, F. (Eds.), *Handbook of Income Distribution Vol.2*. Ámsterdam: Elsevier.
- Álvarez, F. (2019). *Repercusión del Estado de Bienestar. El caso concreto de la sanidad en Aragón* [Trabajo de Fin de Grado]. Universidad de Zaragoza.
- Anghel, B., Basso, H., Bover, O., Casado, J.M., Hospido, L., Izquierdo, M., Kataryniuk, I.A., Lacuesta, A., Montero, J.M., Vozmediano, E. (2018). La desigualdad de la renta, el consumo y la riqueza en España, *Documentos Opcionales* 1806.
- Araoz, O. (2024). *Comparación del impacto en las políticas redistributivas en la desigualdad de la renta en Aragón antes y durante la pandemia: Efectos en los periodos 2019 y 2020 sobre la desigualdad tras la aplicación de políticas redistributivas* [Trabajo de Fin de Grado]. Universidad de Zaragoza.
- Aspachs, O., Durante, R., Graziano, A., Mestres, J., Reynal-Querol, M. y Montalvo J.G. (2021). Tracking the impact of COVID-19 on economic inequality at high frequency, *PLoS ONE* 16(3): e0249121.
- Atkinson, A.B. (1970). On the measurement of Inequality, *Journal of Economic Theory* 2(3) 244-263.
- Atkinson, A.B., Rainwater, L. y Smeeding, T.M. (1995). *Income Distribution in OECD Countries. Evidence from the Luxembourg Income Study*. París: OECD Publishing.
- Ayala, L. (2016). La Desigualdad en España: Fuentes, Tendencias y Comparaciones Internacionales, *Estudios sobre la Economía Española FEDEA 2016/24*.
- Balestra, C. y Oehler, F. (2023). *Measuring the joint distribution of household income, consumption and wealth at the micro level. Methodological issues and experimental results*. Luxemburgo: Publications Office of the European Union.
- Bandrés Moliné, E. (2023). Dimensiones de la desigualdad: España en contexto, *Panorama Social* 37: 53-66.
- Bautista, S. (2020). *Efectos redistributivos del sistema educativo en Aragón durante el periodo 2007-2017: Los efectos en el corto plazo de la educación pública en la*

redistribución de la renta desde la Gran Recesión [Trabajo de Fin de Máster]. Universidad de Zaragoza.

Blumenstock, J., Cadamuro, G., y On, R. (2015). Predicting Poverty and Wealth from Mobile Phone Metadata, *Science* 350(6264): 1073.

Buhmann, B., Rainwater, L., Schmaus, G. y Smeeding, T.M. (1988). Equivalence scales, well-being, inequality, and poverty: sensitivity estimates across ten countries using the Luxembourg Income Study (LIS) database, *Review of Income and Wealth* 34: 115-142.

Cobham, A. y Sumner, A. (2013). Is it all about the Tails? The Palma Measure of Income Inequality, *Centre for Global Development* Working Paper 343.

Coulter, P.B (1989). *Measuring inequality: a methodological handbook*. Boulder: Westview Press.

Cowell, F.A. (2000). Measurement of Inequality. En Atkinson y Bourguignon (Eds.), *Handbook of Income Distribution*. Amsterdam: Elsevier, pp. 1-86.

De Battisti, F., Porro, F. Vernizzi, A. (2019). The Gini coefficient and the case of negative values, *Electronic Journal of Applied Statistical Analysis* 23: 85-107.

Eurostat (2023). *EU statistics on income and living conditions (EU-SILC) methodology – distribution of income*. [https://ec.europa.eu/eurostat/statistics-explained/index.php?title=EU_statistics_on_income_and_living_conditions_\(EU-SILC\)_methodology_-_distribution_of_income#Calculation_method](https://ec.europa.eu/eurostat/statistics-explained/index.php?title=EU_statistics_on_income_and_living_conditions_(EU-SILC)_methodology_-_distribution_of_income#Calculation_method).

FMI (Fondo Monetario Internacional) (2014). Fiscal Policy and Income Inequality, *Policy Papers* 14.

Friedman, M. (1962). *Capitalism and Freedom*, Chicago: University of Chicago Press.

Gini, C. (1912). *Variabilità e mutabilità: contributo allo studio delle distribuzioni e delle relazioni statistiche*, Bolonia: Tipografia di Paolo Cuppini.

Gini, C. (1914). Sulla misura della concentrazione e della variabilità dei caratteri, *Rivista Internazionale di Scienze Sociali e Discipline Ausiliarie* 65.

Gini, C. (2005). On the measurement of concentration and variability of characters, *Metron – International Journal of Statistics* 0: 1:38.

- Goerlich, F.J. (2016). *Distribución de la renta, crisis económica y políticas redistributivas*, Madrid: Fundación BBVA.
- Goerlich, F.J. (2018). Las mil caras de la desigualdad y una más: La obtención de los ingresos en la Encuesta de Condiciones de Vida, *IVIE*.
- Goerlich, F.J. (2020). La Encuesta de Condiciones de Vida: evaluación de los cambios metodológicos en relación a la obtención de los ingresos, *Hacienda Pública Española* 233: 85-116.
- Goerlich, F.J. y Villar, A. (2009). *Desigualdad y Bienestar Social. De la teoría a la práctica*. Bilbao: Fundación BBVA.
- Hagenaars, A., De Vos, K. y Asghar Zaidi, M. (1994). *Poverty Statistics in the Late 1980s: Research Based on Micro-data*. Luxemburgo: Office for Official Publications of the European Communities.
- Hagerbaumer, J.B. (1937). The Gini concentration ratio and the minor concentration ratio: a two parameter index of inequality, *Review of Economics and Statistics* 59.
- INE (Instituto Nacional de Estadística) (2021a). *Encuesta de condiciones de vida (ECV). Ficheros transversales de usuario de la encuesta de 2020*. Madrid: INE.
- INE (Instituto Nacional de Estadística) (2021b). *Encuesta de condiciones de vida (ECV). Base 2013 – Año 2020*. Madrid: INE.
- INE (Instituto Nacional de Estadística) (2023). *Encuesta de condiciones de vida. Metodología*. Madrid: INE.
- INE (Instituto Nacional de Estadística) (2024). *Atlas de Distribución de Renta de los Hogares. Metodología*. Madrid: INE.
- Jean, N., Burke, M., Xie, M., Davis, W. M., Lobell, D. B. y Ermon, S. (2016). Combining Satellite Imagery and Machine Learning to Predict Poverty, *Science* 353(6301): 790.
- Jenkins, S.P. y Lambert P.J. (1993). Ranking income distributions when needs differ, *Review of Income and Wealth* 39: 337-356.
- Joumard, I., Pisu, M., y Bloch, D. (2013). Tackling income inequality: The role of taxes and transfers, *OECD Journal: Economic Studies* 2012(1): 37-70.

Kakwani, N.C. (1977). Measurement of Tax Progressivity: An International Comparison, *Economic Journal* 87: 71-10.

Ley 12/1989, de 9 de mayo, de la Función Estadística Pública. *Boletín Oficial del Estado*, 114, de 11 de mayo de 1989.

López Laborda, J., Marín González, C. y Onrubia, J. (2023). *Observatorio sobre el reparto de los impuestos y las prestaciones entre los hogares españoles. Séptimo informe – 2019 y 2020*. Madrid: Fedea.

Lorenz, M.O. (1905). Methods for measuring the concentration of wealth, *Publications of the American Statistical Association* 9(70): 209-219.

McGregor, T., Smith, B. y Wills, S. (2019). Measuring inequality, *Oxford Review of Economic Policy* 35(3): 368-395.

Musgrave, R.A. (1939). The Voluntary Exchange Theory of Public Economy, *The Quarterly Journal of Economics* 53(2): 213-237.

Naciones Unidas (2024). *Objetivo 10: Reducir la desigualdad en y entre los países*. <https://www.un.org/sustainabledevelopment/es/inequality/>.

OCDE (Organización para la Cooperación y el Desarrollo Económico) (2015). Terms of Reference, OCDE Project on the distribution of household incomes, 2015/2016 Collection.

Ostry, J.D., Berg, A. y Tsangarides, C.G. (2014). *Redistribution, Inequality and Growth*. IMF Staff Discussion Note SDN/14/02.

R Core Team (2023). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Viena, Austria. <https://www.R-project.org/>.

Rawls, J (1971). *A Theory of Justice*. Cambridge: Harvard University Press.

Real Decreto 463/2020, de 14 de marzo, por el que se declara el estado de alarma para la gestión de la situación de crisis sanitaria ocasionada por el COVID-19. *Boletín Oficial del Estado*, 67, de 14 de marzo de 2020.

Real Decreto 1110/2020, de 15 de diciembre, por el que se aprueba el Plan Estadístico Nacional 2021-2024. *Boletín Oficial del Estado*, 340, de 30 de diciembre de 2020.

- Reynolds, M. y Smolensky, E. (1977). *Public Expenditure, Taxes and the Distribution of Income: The United States, 1959, 1961, 1970*. Nueva York: Academic Press.
- Signorell, A. (2024). *DescTools: Tools for Descriptive Statistics*. R package version 0.99.54, <https://CRAN.R-project.org/package=DescTools>.
- Sitthiyot, T. y Holasut, K. (2020). A simple method for measuring inequality, *Palgrave Communications* 6, 112.
- Sohn, A. (2016). *acid: Analysing Conditional Income Distributions*. R package version 1.1, <https://CRAN.R-project.org/package=acid>.
- Stiglitz, J. (2012). *The Price of Inequality: How Today's Divided Society Endangers Our Future*, Nueva York: W. W. Norton & Company.
- Wade, L. (2020). An unequal blow, *Science* 368(6492): 700-703.

Anexos

Anexo I. Código empleado

Dos análisis sobre la función redistributiva del sector público en España: diferencias territoriales y capacidad de reacción ante la pandemia COVID-19

Máster Universitario en Economía - Universidad de Zaragoza

Pablo Padilla Borrajo - Verano 2024

El objetivo del presente código es tratar la Encuesta de Condiciones de Vida para realizar los análisis establecidos en el trabajo. En este sentido, primero se cargan los datos de la Encuesta de Condiciones de Vida para poder trabajar con ellos. Asimismo, se imputan las rentas de acuerdo con la metodología presentada en el trabajo. Se calculan el índice de Gini, el índice de Reynolds-Smolensky y la media de la renta por tramos. Finalmente, se realizan algunos estudios complementarios.

Se selecciona el directorio y se cargan las librerías

```
library(dplyr)
library(sqldf)
library(DescTools)
library(acid)
library(survey)
```

En primer lugar, se crea una tabla paramétrica asignando cada uno de los códigos asignados a las comunidades autónomas por parte de la ECV (campo DB040) con los nombres oficiales de las distintas regiones. Esto será usado más adelante en el código

```
# Creo una tabla paramétrica con las comunidades autónomas
Tabla_param_comunidades <- data.frame(
  DB040 = c("ES11", "ES12", "ES13", "ES21", "ES22", "ES23", "ES24", "ES30", "ES41", "ES42", "ES43", "ES51", "ES52", "ES53", "ES61", "ES62", "ES63", "ES64", "ES70"),
  CCAA = c("Galicia", "Principado de Asturias", "Cantabria", "País Vasco", "Comunidad Foral de Navarra", "La Rioja", "Aragón", "Comunidad de Madrid", "Castilla y León", "Castilla-La Mancha", "Extremadura", "Cataluña", "Comunidad Valenciana", "Illes Balears", "Andalucía", "Región de Murcia", "Ciudad Autónoma de Ceuta", "Ciudad Autónoma de Melilla", "Canarias"))
```

Se crea una función llamada tratamiento_datos de la que obtengo una lista con dos dataframes, uno para individuos y otro para hogares con todo los datos que se van a utilizar posteriormente

```
tratamiento_datos <- function(Micro_Td,Meta_Td,Micro_Th,Meta_Th,Micro_Tp,Meta_Tp,Micro_Tr,Meta_Tr) {
  # Se guarda e imprime la fecha en la que se inicia la función
  start.time <- Sys.time()
  cat("\n")
  cat("\n Inicio del proceso de carga de datos: ")
  print.Date(start.time)
  t0 <- proc.time()

  # Se crea una función llamada carga_datos para cargar los ficheros de la ECV obtenidos del INE. Esta función está adaptada de la proporcionada directamente por el INE
  carga_datos <- function(fichero_micro,fichero_meta) {

    # Lectura del fichero de metadatos (METAD), Hoja "Diseño" de archivo .xlsx
    tryCatch((workBook <- loadWorkbook(fichero_meta)), error=function(e)
      stop(paste("Error. No se puede abrir el fichero: ", e, fichero_meta, ". Saliendo de la ejecución...", sep = "")))
    df <- readNamedRegion(workBook, name = "METADATOS")

    nombresVarbls <- df[,1]
    nombresTablas <- df[,2]
    posiciones <- df[,3]
    tipo <- df[,4]
    tamano <- length(nombresVarbls)

    # Lectura del fichero de microdatos (MICROD)
    if(length(df) == 4){

      # Capturamos las columnas con su tipo de dato
      tipDatos <- as.vector(sapply(df[,4], function(x){
        if(identical(x, "A"))
          "character"
        else{
          if(identical(x, "N"))
            "numeric"
        }
      }))
      # Lectura del fichero de microdatos (MICROD), decimales con punto en MD
      tryCatch((df1 <- read.fwf(file = fichero_micro, widths= posiciones, colClasses=tipDatos)), error=function(e)
        stop(paste("Error. No se encuentra el fichero: ", e, fichero_micro, ". Saliendo de la ejecución...", sep = "")))

    }else{
      formatos <- df[,5]

      # Lectura del fichero de microdatos (MICROD), decimales sin punto en MD
    }
  }
}
```

```

tryCatch((df1 <- read.fortran(file = fichero_micro, format= formatos)), error=function(e)
  stop(paste("Error. No se encuentra el fichero: ", e, fichero_micro, ". Saliendo de la ejecución...", sep = "")))
}

#Aplicamos los nombres de la cabecera del registro
names(df1) <- df[,1]

return(df1)
}

# En total disponemos de 4 archivos distintos
# TD: fichero de datos básicos del hogar
# TR: fichero de datos básicos de la persona
# TP: fichero de datos detallados de los adultos.
# TH: fichero de datos detallados del hogar

Tp <- carga_datos(Micro_Tp,Meta_Tp)
cat("\n Tp Cargado")
Tr <- carga_datos(Micro_Tr,Meta_Tr)
cat("\n Tr Cargado")
Th <- carga_datos(Micro_Th,Meta_Th)
cat("\n Th Cargado")
Td <- carga_datos(Micro_Td,Meta_Td)
cat("\n Td Cargado")

#####
# Una vez disponemos de los datos cargados en R ya se puede proceder a realizar los tratamientos
#####

# Creamos dos df: uno para hogares y otro para cada uno de los individuos, para ello unimos los ficheros que tienen información a ese nivel

individuos <- sqldf("
  SELECT a.*, b.*
  FROM Tr AS a
    LEFT JOIN Tp AS b
    ON a.RB030=b.PB030
  ")

hogares <- sqldf("
  SELECT a.*, b.*
  FROM Th AS a
    LEFT JOIN Td AS b
    ON a.HB030=b.DB030
  ")

# Calculo el indicador de hogar
individuos$id_hogar <- floor(individuos$RB030 / 100)

#####
# Cálculo de Variables a nivel individuo #####
# Rentas salariales de los individuos
individuos$Rentas_sala_indiv <- with(individuos,PY010G+PY021G)

# Rentas de autónomos
individuos$Rentas_auto_indiv <- with(individuos,PY050G)

# Rentas del capital
individuos$Rentas_capital_indiv <- with(individuos,PY080G)

# Rentas de mercado (agregación de las anteriores más cotizaciones sociales a cargo del empleador)
individuos$Rentas_mdo_indiv <- with(individuos,Rentas_sala_indiv+Rentas_auto_indiv+Rentas_capital_indiv+PY030G)

# Transferencias
individuos$Transferencias_indiv <- with(individuos,PY090G+PY100G+PY110G+PY120G+PY130G+PY140G)

# Se imprime que las variables a nivel individual ya están calculadas
cat("\n Variables a nivel individuo calculadas ")

# Una vez tenemos las variables calculadas a nivel de individuo, las agregamos por hogar en la tabla de hogares
hogares <- hogares %>%
  left_join(
    individuos %>%
      group_by(id_hogar) %>%
      summarise(
        Rentas_mdo_indiv_agr_hogar = sum(Rentas_mdo_indiv, na.rm = TRUE),
        Transferencias_indiv_agr_hogar = sum(Transferencias_indiv, na.rm = TRUE),
        Cotiz_empleador_indiv_agr_hogar = sum(PY030G, na.rm = TRUE)
      ),
    by = c("DB030" = "id_hogar")
  )

#####
# Cálculo de Variables a nivel hogar #####
# Renta de Mercado
hogares$Renta_mdo_hogar <- with(hogares,Rentas_mdo_indiv_agr_hogar+HY090G+HY040G)

```

```

# Renta Privada
hogares$Renta_priv_hogar <- with(hogares,Renta_mdo_hogar+HY080G-HY130G)

# Transferencias
hogares$Transferencias_hogar <- with(hogares,Transferencias_indiv_agr_hogar+HY050G+HY060G+HY070G+HY080G+HY110G)

# Renta Bruta
hogares$Renta_bruta_hogar <- with(hogares,Renta_mdo_hogar+Transferencias_hogar)

# Renta Privada con las transferencias públicas
hogares$Renta_priv_transf_hogar <- with(hogares,Renta_bruta_hogar-HY130G)

# Impuestos
hogares$Impuestos_hogar <- with(hogares,HY120G+HY140G)

# Renta Disponible
hogares$Renta_disponible_hogar <- with(hogares,Renta_bruta_hogar-Impuestos_hogar-HY130G-Cotiz_empleador_indiv_agr_hogar)

# Se imprime que las variables a nivel hogar ya están calculadas
cat("\n Variables a nivel hogar calculadas")

# Comprobamos que la variable renta disponible calculada en el código coincide con la variable vhRentaa calculada por el INE
chequeo <- sqldf(" SELECT DB030, DB040, Renta_disponible_hogar, vhRentaa,
                    Renta_disponible_hogar - vhRentaa AS Dif_vhRentaa
                FROM hogares
                WHERE ABS(Renta_disponible_hogar - vhRentaa)>0.1
            ")

# Si hay un hogar donde no coincide la renta disponible calculada con la del INE se imprime dicho mensaje
if (nrow(chequeo) > 0) {
  cat("\n Diferencias relevantes en la renta disponible")
  print(chequeo)
} else {
  cat("\n Sin diferencias relevantes en la renta disponible")
}

# Aplicamos las escalas de equivalencia para todas las variables de renta, concretamente se aplican las de la OCDE que no resulta necesario calcularlas ya que vienen informadas en las tablas de hogar
hogares$Renta_mdo_equi <- with(hogares,Renta_mdo_hogar/HX240)

hogares$Renta_priv_equi <- with(hogares,Renta_priv_hogar/HX240)
hogares$Renta_bruta_equi <- with(hogares,Renta_bruta_hogar/HX240)
hogares$Renta_priv_transf_equi <- with(hogares,Renta_priv_transf_hogar/HX240)
hogares$Renta_disponible_equi <- with(hogares,Renta_disponible_hogar/HX240)

# Cruzamos la tabla de hogares con la tabla paramétrica de las CCAA para tener no solo el código sino también el nombre
hogares_final <- sqldf("
    SELECT a.*, b.CCAA
    FROM hogares AS a
    LEFT JOIN Tabla_param_comunidades AS b
    ON a.DB040=b.DB040
")

# Pegamos en la tabla de individuos todas las variables que hemos calculado a nivel hogar previamente
individuos_final <- sqldf("
    SELECT a.*, b.*
    FROM individuos AS a
    LEFT JOIN hogares_final AS b
    ON a.id_hogar=b.DB030
")

# Se imprime el mensaje final y el tiempo que ha tardado la función en ejecutarse
end.time <- Sys.time()
cat("\n Fin del proceso de carga de datos: ")
print.Date(end.time)
TTotal <- proc.time() - t0
tiempo <- TTotal[3]
formato_tiempo <- function(tiempo) {
  horas <- floor(tiempo / 3600)
  minutos <- floor((tiempo %% 3600) / 60)
  segundos <- round(tiempo %% 60, 0)
  partes <- c(
    if (horas > 0) paste(horas, ifelse(horas == 1, "hora", "horas")),
    if (minutos > 0) paste(minutos, ifelse(minutos == 1, "minuto", "minutos")),
    if (segundos > 0) paste(segundos, ifelse(segundos == 1, "segundo", "segundos"))
  )
  return(paste(partes, collapse = " y "))
}
cat("\n Tiempo transcurrido:", formato_tiempo(tiempo))

# Hacemos que la función devuelva una lista con dos df, hogares e individuos

```

```
return(list(hogares = hogares_final, individuos = individuos_final))
}
```

Se ejecuta la función anterior para las ECV de los años 2020 y 2021. Hay que tener en cuenta que las ECV contienen información relativa al año anterior. Así, la ECV del año 2020 nos da información sobre las rentas de los individuos en el año 2019 y la ECV del 2021 proporciona información acerca de las rentas obtenidas en el año 2021. La fuente de la que se obtienen los ficheros originales es el INE a través de la siguiente dirección: https://www.ine.es/dyngs/INEbase/es/operacion.htm?c=Estadistica_C&cid=1254736176807&menu=resultados&idp=1254735976608#_tabs-1254736195153

La función tiene la siguiente sintaxis:

```
# tratamiento_datos(Micro_Td,Meta_Td, Micro_Th,Meta_Th, Micro_Tp,Meta_Tp, Micro_Tr,Meta_Tr)
```

#Ejecutamos para La ECV 2020

```
datos_2020 <-tratamiento_datos("md_ECV_Td_2020.txt","dr_ECV_Td_2016a2020.xlsx","md_ECV_Th_2020.txt","dr_ECV_Th_2020.xlsx","md_ECV_Tp_2020.txt","dr_ECV_Tp_2020.xlsx","md_ECV_Tr_2020.txt","dr_ECV_Tr_2016a2020.xlsx")
```

Separamos Los dos df

```
hogares_2020 <- datos_2020[[1]]
individuos_2020 <- datos_2020[[2]]
```

#Ejecutamos para La ECV 2021

```
datos_2021 <-tratamiento_datos("md_ECV_Td_2021.txt","dr_ECV_Td_2021.xlsx","md_ECV_Th_2021.txt","dr_ECV_Th_2021.xlsx","md_ECV_Tp_2021.txt","dr_ECV_Tp_2021.xlsx","md_ECV_Tr_2021.txt","dr_ECV_Tr_2021.xlsx")
```

Separamos Los dos df

```
hogares_2021 <- datos_2021[[1]]
individuos_2021 <- datos_2021[[2]]
```

Se crea una función para realizar los cálculos del índice de Gini

```
calculo_gini <- function(table_in) {
```

Se guarda e imprime La fecha en la que se inicia La función

```
start.time <- Sys.time()
```

```
cat("\n")
```

```
cat("\n Inicio del proceso de cálculo de Gini: ")
```

```
print.Date(start.time)
```

```
t0 <- proc.time()
```

Creamos un dataframe donde van a estar Los resultados obtenidos

```
resultados <- data.frame(CCAA = c(Tabla_param_comunidades$CCAA, "España"),
                           gini_mdo = NA,
                           gini_priv = NA,
                           gini_bruta = NA,
                           gini_disponible = NA)
```

Calculamos Los Ginis para las variables de renta para el conjunto de La población en España. Para ello utilizamos La función weighted.gini

```
resultados$gini_mdo[resultados$CCAA == "España"] <- weighted.gini(table_in$Renta_mdo_equi, table_in$RB050)[[1]]
```

```
resultados$gini_priv[resultados$CCAA == "España"] <- weighted.gini(table_in$Renta_priv_equi, table_in$RB050)[[1]]
```

```
resultados$gini_bruta[resultados$CCAA == "España"] <- weighted.gini(table_in$Renta_bruta_equi, table_in$RB050)[[1]]
```

```
resultados$gini_disponible[resultados$CCAA == "España"] <- weighted.gini(table_in$Renta_disponible_equi, table_in$RB050)[[1]]
```

Calculamos Los Ginis para las variables de renta para cada una de las CCAA. Se realiza a través de un bucle recorriendo por completo La tabla paramétrica

```
for (CCAA in Tabla_param_comunidades$CCAA) {
  sub_tabla <- table_in[table_in$CCAA == CCAA, ]
```

```
resultados$gini_mdo[resultados$CCAA == CCAA] <- weighted.gini(sub_tabla$Renta_mdo_equi, sub_tabla$RB050)[[1]]
```

```
resultados$gini_priv[resultados$CCAA == CCAA] <- weighted.gini(sub_tabla$Renta_priv_equi, sub_tabla$RB050)[[1]]
```

```
resultados$gini_bruta[resultados$CCAA == CCAA] <- weighted.gini(sub_tabla$Renta_bruta_equi, sub_tabla$RB050)[[1]]
```

```
resultados$gini_disponible[resultados$CCAA == CCAA] <- weighted.gini(sub_tabla$Renta_disponible_equi, sub_tabla$RB050)[[1]]
```

```
}
```

Se imprime el mensaje final y el tiempo que ha tardado La función en ejecutarse

```
end.time <- Sys.time()
```

```
cat("\n")
```

```
cat("\n Fin del proceso de cálculo de Gini: ")
```

```
print.Date(end.time)
```

```
TTotal <- proc.time() - t0
```

```
tiempo <- TTotal[3]
```

```
formato_tiempo <- function(tiempo) {
```

```
  horas <- floor(tiempo / 3600)
```

```
  minutos <- floor((tiempo %% 3600) / 60)
```

```
  segundos <- round(tiempo %% 60, 0)
```

```
  partes <- c(
```

```
    if (horas > 0) paste(horas, ifelse(horas == 1, "hora", "horas")),
```

```

    if (minutos > 0) paste(minutos, ifelse(minutos == 1, "minuto", "minutos")),
    if (segundos > 0) paste(segundos, ifelse(segundos == 1, "segundo", "segundos"))
  )
return(paste(partes, collapse = " y "))
}
cat("\n Tiempo transcurrido:", formato_tiempo(tiempo))

# Eliminamos las columnas no utilizadas y devolvemos el resultado
resultados <- resultados[ , colSums(is.na(resultados)) < nrow(resultados)]
return(resultados)
}

Se ejecuta la función anterior para los datos obtenidos de las ECV de los años 2020 y 2021.

# La función tiene la siguiente sintaxis:
# calculo_gini(table_in)

# Calculamos los Ginis para los datos de la ECV del año 2020
gini_individuos_2020 <- calculo_gini(individuos_2020)

# Calculamos los Ginis para los datos de la ECV del año 2021
gini_individuos_2021 <- calculo_gini(individuos_2021)

Se crea una función para realizar los cálculos del índice de Reynolds-Smolensky

reynolds_smolensky <- function(table_in) {

# Se guarda e imprime la fecha en la que se inicia la función
start.time <- Sys.time()
cat("\n")
cat("\n Inicio del proceso de cálculo de Reynolds-Smolensky: ")
print.Date(start.time)
t0 <- proc.time()

# Creamos un dataframe donde van a estar los resultados obtenidos
resultados <- data.frame(CCAA = c(Tabla_param_comunidades$CCAA, "España"),
                           gini_priv = NA,
                           conc_priv_trans = NA,
                           conc_disponible = NA,
                           rs_trans = NA,
                           rs_imp = NA,
                           rs_tot = NA)

# Creamos una función para calcular los índices de concentración. Para ello nos basamos en lo que se está utilizando cuando se ejecuta la función weighted.gini
concentracion <- function (x, w)
{
  n <- length(x)
  w <- w/sum(w)
  w.c <- cumsum(w)
  xw.c <- cumsum(w * x)
  xw.c <- xw.c/xw.c[n]
  Gini <- t(xw.c[-1]) %*% w.c[-n] - t(xw.c[-n]) %*% w.c[-1]
  list(Gini = Gini)
}

# Calculamos los índices de concentración para el conjunto de la población de España ordenando a los individuos en función de la Renta privada equivalente, el punto de partida del análisis. El índice de concentración de la renta privada equivalente será igual a su índice de Gini
resultados$gini_priv[resultados$CCAA == "España"] <- concentracion(table_in$Renta_priv_equi[order(table_in$Renta_priv_equi)],
table_in$RB050[order(table_in$Renta_priv_equi)])

resultados$conc_priv_trans[resultados$CCAA == "España"] <- concentracion(table_in$Renta_priv_transf_equi[order(table_in$Renta_priv_equi)],
table_in$RB050[order(table_in$Renta_priv_equi)])

resultados$conc_disponible[resultados$CCAA == "España"] <- concentracion(table_in$Renta_disponible_equi[order(table_in$Renta_priv_equi)],
table_in$RB050[order(table_in$Renta_priv_equi)])

# Obtenemos los índices de Reynolds-Smolensky para el conjunto de España
resultados$rs_trans[resultados$CCAA == "España"] <- (resultados$gini_priv[resultados$CCAA == "España"][[1]][1, 1] - resultados$conc_priv_trans[resultados$CCAA == "España"][[1]][1, 1])

resultados$rs_imp[resultados$CCAA == "España"] <- (resultados$conc_priv_trans[resultados$CCAA == "España"][[1]][1, 1] - resultados$conc_disponible[resultados$CCAA == "España"][[1]][1, 1])

resultados$rs_tot[resultados$CCAA == "España"] <- (resultados$gini_priv[resultados$CCAA == "España"][[1]][1, 1] - resultados$conc_disponible[resultados$CCAA == "España"][[1]][1, 1])

# Calculamos los índices de concentración y de Reynolds-Smolensky para cada una de las CCAA. Se realiza a través de un bucle recorriendo por completo la tabla paramétrica
for (CCAA in Tabla_param_comunidades$CCAA) {
  sub_tabla <- table_in[table_in$CCAA == CCAA, ]

  resultados$gini_priv[resultados$CCAA == CCAA] <- concentracion(sub_tabla$Renta_priv_equi[order(sub_tabla$Renta_priv_equi)],
sub_tabla$RB050[order(sub_tabla$Renta_priv_equi)])
}

```

```

    resultados$conc_priv_trans[resultados$CCAA == CCAA] <- concentracion(sub_tabla$Renta_priv_transf_equi[order(sub_tabla$Renta_priv_equi)], sub_tabla$RB050[order(sub_tabla$Renta_priv_equi)])}

    resultados$conc_disponible[resultados$CCAA == CCAA] <- concentracion(sub_tabla$Renta_disponible_equi[order(sub_tabla$Renta_priv_equi)], sub_tabla$RB050[order(sub_tabla$Renta_priv_equi)])}

    resultados$rs_trans[resultados$CCAA == CCAA] <- (resultados$gini_priv[resultados$CCAA == CCAA][[1]][1, 1] - resultados$conc_priv_trans[resultados$CCAA == CCAA][[1]][1, 1])

    resultados$rs_imp[resultados$CCAA == CCAA] <- (resultados$conc_priv_trans[resultados$CCAA == CCAA][[1]][1, 1] - resultados$conc_disponible[resultados$CCAA == CCAA][[1]][1, 1])

    resultados$rs_tot[resultados$CCAA == CCAA] <- (resultados$gini_priv[resultados$CCAA == CCAA][[1]][1, 1] - resultados$conc_disponible[resultados$CCAA == CCAA][[1]][1, 1])
}

# Se imprime el mensaje final y el tiempo que ha tardado la función en ejecutarse
end.time <- Sys.time()
cat("\n")
cat("\n Fin del proceso de cálculo de Reynolds-Smolensky: ")
print.Date(end.time)
TTotal <- proc.time() - t0
tiempo <- TTotal[3]
formato_tiempo <- function(tiempo) {
  horas <- floor(tiempo / 3600)
  minutos <- floor((tiempo %% 3600) / 60)
  segundos <- round(tiempo %% 60, 0)

  partes <- c(
    if (horas > 0) paste(horas, ifelse(horas == 1, "hora", "horas")),
    if (minutos > 0) paste(minutos, ifelse(minutos == 1, "minuto", "minutos")),
    if (segundos > 0) paste(segundos, ifelse(segundos == 1, "segundo", "segundos"))
  )
  return(paste(partes, collapse = " y "))
}
cat("\n Tiempo transcurrido:", formato_tiempo(tiempo))

# Eliminamos las columnas no utilizadas y devolvemos el resultado
resultados <- resultados[, colSums(is.na(resultados)) < nrow(resultados)]
return(resultados)
}

```

Se ejecuta la función anterior para los datos obtenidos de las ECV de los años 2020 y 2021.

```

# La función tiene la siguiente sintaxis:
# reynolds_smolensky(table_in)

# Calculamos los Reynolds-Smolensky para los datos de la ECV del año 2020
reynolds_2020 <- reynolds_smolensky(individuos_2020)
# Calculamos los Reynolds-Smolensky para los datos de la ECV del año 2021
reynolds_2021 <- reynolds_smolensky(individuos_2021)

```

Se crea una función para realizar los cálculos de las medias poblacionales por tramos de renta bruta

```

tramos_de_renta <- function(table_in) {
  # Se guarda e imprime la fecha en la que se inicia la función
  start.time <- Sys.time()
  cat("\n")
  cat("\n Inicio del proceso de media por tramos de renta: ")
  print.Date(start.time)
  t0 <- proc.time()

  # Creamos un dataframe auxiliar que se va a ir utilizando a lo largo de la función
  table_aux <- table_in %>% select(RB050, Renta_mdo_equi, Renta_priv_equi, Renta_bruta_equi, Renta_disponible_equi)

  # Creamos un diseño de encuesta ponderado
  design <- svydesign(ids = ~1, data = table_aux, weights = ~RB050)

  # Calculamos los percentiles ponderados para la variable Renta Bruta
  percentiles_list <- svyquantile(~Renta_bruta_equi, design, quantiles = seq(0, 1, by = 0.01))

  # Convertimos la lista de percentiles en un vector numérico y eliminamos los valores NaN
  percentiles_vector <- unlist(percentiles_list)
  percentiles_vector <- na.omit(percentiles_vector)

  # Se crean los grupos de quintiles, deciles y centiles basados en los percentiles ponderados
  table_aux <- table_aux %>%
    mutate(
      Quintil = cut(Renta_bruta_equi, breaks = quantile(Renta_bruta_equi, probs = seq(0, 1, by = 0.2), na.rm = TRUE),
                    labels = 1:5, include.lowest = TRUE),
      Decil = cut(Renta_bruta_equi, breaks = quantile(Renta_bruta_equi, probs = seq(0, 1, by = 0.1), na.rm = TRUE),
                 labels = 1:10, include.lowest = TRUE),
      Centil = cut(Renta_bruta_equi, breaks = quantile(Renta_bruta_equi, probs = seq(0, 1, by = 0.01), na.rm = TRUE),
                   labels = 1:100, include.lowest = TRUE)
    )
}

```

```

)
# Hacemos una comprobación para ver si los intervalos son coincidentes
if (length(unique(table_aux$Quintil)) != 5 || length(unique(table_aux$Decil)) != 10 || length(unique(table_aux$Centil)) != 100)
{
  stop("El número de etiquetas no coincide con el número de intervalos.")
}

# Actualizamos el diseño de la encuesta con los grupos creados
design <- update(design, Quintil = table_aux$Quintil, Decil = table_aux$Decil, Centil = table_aux$Centil)

# Calculamos las medias ponderadas por quintil
quintiles <- svyby(~Renta_mdo_equi + Renta_priv_equi + Renta_bruta_equi + Renta_disponible_equi, by = ~Quintil, design, svymean)
n)

# Calculamos las medias ponderadas por decil
deciles <- svyby(~Renta_mdo_equi + Renta_priv_equi + Renta_bruta_equi + Renta_disponible_equi, by = ~Decil, design, svymean)

# Calculamos las medias ponderadas para los centiles 91-99 y el centil 100
centiles <- svyby(~Renta_mdo_equi + Renta_priv_equi + Renta_bruta_equi + Renta_disponible_equi, by = ~Centil, design, svymean)
%>%
filter(Centil %in% c(91:99, 100)) %>%
group_by(Centil = ifelse(Centil == 100, "Centil 100", "Centil 91-99")) %>%
summarise(across(starts_with("renta_"), mean, na.rm = TRUE))

# Unimos los resultados obtenidos para los tramos en una tabla auxiliar
tabla_aux <- bind_rows(
  quintiles %>% filter(Quintil %in% 1:4) %>% mutate(Grupo = paste("Quintil", Quintil)),
  deciles %>% filter(Decil == 9) %>% mutate(Grupo = "Decil 9"),
  centiles %>% rename(Grupo = Centil)
)

# Calculamos una fila con los totales
totales <- svymean(~Renta_mdo_equi + Renta_priv_equi + Renta_bruta_equi + Renta_disponible_equi, design) %>%
  as.data.frame() %>%
  t() %>%
  as.data.frame() %>%
  setNames(c("Renta_mdo_equi", "Renta_priv_equi", "Renta_bruta_equi", "Renta_disponible_equi")) %>%
  mutate(Grupo = "Total")

# Combinamos en una tabla final la tabla auxiliar con la fila de totales
resultados <- bind_rows(tabla_aux %>% select(Grupo, Renta_mdo_equi, Renta_priv_equi, Renta_bruta_equi, Renta_disponible_equi),
  totales %>% slice(1) %>% select(Grupo, Renta_mdo_equi, Renta_priv_equi, Renta_bruta_equi, Renta_disponible_equi))

# Se imprime el mensaje final y el tiempo que ha tardado la función en ejecutarse
end.time <- Sys.time()
cat("\n")
cat("\n Fin del proceso de cálculo de media por tramos de renta: ")
print.Date(end.time)
TTotal <- proc.time() - t0
tiempo <- TTotal[3]
formato_tiempo <- function(tiempo) {
  horas <- floor(tiempo / 3600)
  minutos <- floor((tiempo %% 3600) / 60)
  segundos <- round(tiempo %% 60, 0)

  partes <- c(
    if (horas > 0) paste(horas, ifelse(horas == 1, "hora", "horas")),
    if (minutos > 0) paste(minutos, ifelse(minutos == 1, "minuto", "minutos")),
    if (segundos > 0) paste(segundos, ifelse(segundos == 1, "segundo", "segundos"))
  )
  return(paste(partes, collapse = " y "))
}
cat("\n Tiempo transcurrido:", formato_tiempo(tiempo))

# Devuelvo los resultados obtenidos
return(resultados)
}

```

Se ejecuta la función anterior para los datos obtenidos de las ECV de los años 2020 y 2021.

```

# La función tiene la siguiente sintaxis:
# tramos_de_renta(table_in)

# Calculamos las medias ponderadas de los tramos de renta para los datos de la ECV del año 2020
tramos_renta_2020 <- tramos_de_renta(individuos_2020)

# Calculamos las medias ponderadas de los tramos de renta para los datos de la ECV del año 2020
tramos_renta_2021 <- tramos_de_renta(individuos_2021)

```

De aquí en adelante, se realizan distintas ejecuciones acerca de aspectos menores del trabajo

Calculamos distintas variables

```

# Número de habitantes en España
sum(individuos_2020$RB050)
sum(individuos_2021$RB050)

# Número de habitantes en España
sum(hogares_2020$DB090)
sum(hogares_2021$DB090)

# Transferencias a otros hogares emitidas
sum(hogares_2020$DB090*hogares_2020$HY1306)
sum(hogares_2021$DB090*hogares_2021$HY1306)

# Transferencias recibidas de otros hogares
sum(hogares_2020$DB090*hogares_2020$HY0806)
sum(hogares_2021$DB090*hogares_2021$HY0806)

```

Se realiza un estudio acerca de las rentas negativas

```

# Vamos a ver cuántos individuos tienen una renta negativa por CCAA (1%-3%)
sqldf("SELECT DISTINCT CCAA, COUNT(*) AS TOTAL,
       SUM(CASE WHEN Renta_priv_equi<0
                 THEN 1 ELSE 0 END) AS NEGATIVOS
  FROM (SELECT CCAA,Renta_priv_equi FROM individuos_2020
        UNION ALL SELECT CCAA,Renta_priv_equi FROM individuos_2021)
 GROUP BY CCAA")

# Juntamos Los individuos de Los dos años en una tabla, vamos a hacer el estudio sobre La renta disponible y La renta de merca
do
individuos_20_21 <- sqldf("

  SELECT Renta_disponible_equi,Renta_mdo_equi, RB050
  FROM individuos_2020
  UNION ALL
  SELECT Renta_disponible_equi,Renta_mdo_equi, RB050
  FROM individuos_2021
  ")

# Hacemos distintos tratamientos que propone la literatura
ind_trat <- individuos_20_21 %>%
  mutate(
    Renta_disponible_equi_cero = ifelse(Renta_disponible_equi < 0, 0, Renta_disponible_equi),
    Renta_disponible_equi_na = ifelse(Renta_disponible_equi < 0, NA, Renta_disponible_equi),
    Renta_mdo_equi_cero = ifelse(Renta_mdo_equi < 0, 0, Renta_mdo_equi),
    Renta_mdo_equi_na = ifelse(Renta_mdo_equi < 0, NA, Renta_mdo_equi))

# Calculamos Los índices de Gini para observar que no hay grandes diferencias
# Renta de mercado
weighted.gini(individuos_20_21$Renta_mdo_equi,individuos_20_21$RB050)
weighted.gini(ind_trat$Renta_mdo_equi_cero,ind_trat$RB050)
Gini(ind_trat$Renta_mdo_equi_na,weights = ind_trat$RB050, unbiased = FALSE, na.rm = TRUE)

# Renta disponible
weighted.gini(individuos_20_21$Renta_disponible_equi,individuos_20_21$RB050)
weighted.gini(ind_trat$Renta_disponible_equi_cero,ind_trat$RB050)
Gini(ind_trat$Renta_disponible_equi_na,weights = ind_trat$RB050, unbiased = FALSE, na.rm = TRUE)

```

Se crean distintas tablas para la realización de gráficos

```

# Ilustración 3.1
z3.1_ecs_de_escala <- sqldf("SELECT DISTINCT
                                TAMAÑO, SUM(VhRentaa*DB090)/SUM(DB090) AS Renta_por_hogar,
                                SUM(VhRentaa/HX040*DB090)/SUM(DB090) AS Renta_pc
                               FROM (SELECT HX040, CASE WHEN HX040 >5 THEN 5 ELSE HX040 END AS TAMAÑO, VhRentaa, DB090
                                     FROM hogares_2020
                                     UNION ALL
                                     SELECT HX040, CASE WHEN HX040 >5 THEN 5 ELSE HX040 END AS TAMAÑO, VhRentaa, DB090
                                     FROM hogares_2021)
                                GROUP BY TAMAÑO
                               ")

# Ilustración 3.2
z3.2_mejor_escala_equi <- sqldf("SELECT DISTINCT
                                TAMAÑO, SUM(VhRentaa*DB090)/SUM(DB090)/HX040 AS Renta_pc,
                                SUM((VhRentaa/POWER(HX040,0.5))*DB090)/SUM(DB090) AS Renta_ocde,
                                SUM(VhRentaa/HX240*DB090)/SUM(DB090) AS Renta_ocde_mod
                               FROM (SELECT HX240, HX040, CASE WHEN HX040 >5 THEN 5 ELSE HX040 END AS TAMAÑO, VhRentaa, DB090
                                     FROM hogares_2020
                                     UNION ALL
                                     SELECT HX240, HX040, CASE WHEN HX040 >5 THEN 5 ELSE HX040 END AS TAMAÑO, VhRentaa, DB090
                                     FROM hogares_2021)
                                GROUP BY TAMAÑO
                               ")

```