

PAPELES DE TRABAJO

11/2022

Recaudación fiscal por IRPF y desigualdad salarial en presencia de datos agrupados (*)

ALEXANDRA SOBERÓN (**)

Universidad de Cantabria

(*) Documento elaborado en el marco del proyecto investigación sobre “Recaudación fiscal por IRPF y desigualdad salarial en presencia de datos agrupados. Un enfoque de elasticidades”, financiado por el Instituto de Estudios Fiscales.

(**) La autora agradece la financiación recibida del Instituto de Estudios Fiscales para la realización de este proyecto. Asimismo, agradece especialmente el apoyo prestado por el profesor Antonio de las Heras Pérez sin el cual este proyecto nunca podría haberse llevado a cabo.



ÍNDICE

Resumen

1. INTRODUCCIÓN

2. REVISIÓN DE LA LITERATURA

2.1. Literatura sobre desigualdad salarial

2.2. Literatura sobre desigualdad e ingresos por IRPF

3. DATOS

4. METODOLOGÍA Y PROCEDIMIENTO DE ESTIMACIÓN

4.1. Estimación no paramétrica de la densidad con datos agrupados

4.2. Desigualdad salarial y recaudación fiscal por IRPF

4.2.1. Elasticidad fiscal y distribución de la renta

4.2.2. Elasticidad fiscal y desigualdad salarial

5. RESULTADOS EMPÍRICOS

5.1. Distribución no paramétrica de las rentas salariales

5.2. Elasticidad fiscal a cambios en la distribución de la renta

5.3. Elasticidad fiscal a cambios en la desigualdad salarial

6. CONCLUSIONES

ANEXO

BIBLIOGRAFÍA

Resumen

El contexto de gran incertidumbre internacional que se prevé sea predominante en los próximos años pone en cuestión la sostenibilidad del Estado del Bienestar español. Con el objetivo de analizar la suficiencia financiera del Sistema Fiscal español, en este trabajo se analiza el comportamiento de los ingresos fiscales por IRPF ante variaciones en las rentas salariales, así como en la desigualdad salarial utilizando registros administrativos recogidos por la Agencia Española de la Administración Tributaria (AEAT). Para obtener estimadores consistentes y evitar problemas relacionados con la mala especificación del modelo econométrico, en la primera parte del trabajo se propone una técnica de estimación novedosa que permite estimar consistentemente funciones de densidad de los ingresos de los asalariados cuando los datos se encuentran recogidos en intervalos asimétricos y la variable de interés sólo tiene soporte positivo. Posteriormente, se establece un marco teórico flexible que nos permite estimar consistentemente la elasticidad de la recaudación fiscal por IRPF con respecto a variaciones en los rendimientos salariales y en la desigualdad salarial a través de técnicas no paramétricas. Finalmente, utilizando los datos de la Estadística de Declarantes del IRPF se estiman dichas elasticidades a lo largo del periodo comprendido entre los años 2003 a 2019 y se identifican aquellos grupos poblacionales en los que incrementos en los tipos marginales (o medios) del IRPF realmente implican aumentos en la recaudación fiscal.

1. INTRODUCCIÓN

La desigualdad salarial de una economía constituye un aspecto económico fundamental dado que impacta en el crecimiento económico de los países a través de distintos canales, tal y como se destaca en (Arranz, *et al.*, 2015), (García-Serrano y Arranz, 2013) y (Simón, 2009), entre otros. Por un lado, la dispersión salarial puede influir en la eficiencia del funcionamiento del mercado de trabajo generando incentivos a la inversión en capital humano y, por tanto, al crecimiento económico (Mueller *et al.*, 2017). Por otro lado, la desigualdad en los salarios también interviene en la distribución de la renta y su evolución en el tiempo, influyendo en la desigualdad de los ingresos y, a través de ella, en la cohesión social y la equidad. Asimismo, el nivel de riqueza de cada hogar puede afectar a sus decisiones de consumo e inversión y al efecto esperado que determinadas políticas fiscales y monetarias pueden tener sobre ellos. A este respecto se puede encontrar información adicional en (Grossman, 1991), (Persson *et al.*, 1994), (Alesina y Rodrik, 1994), entre otros.

Durante las últimas décadas se han producido transformaciones profundas en la economía y la sociedad (dentro de las que se encuentran el aumento de la competencia extranjera en los mercados de bienes y servicios, la feminización del mercado de trabajo, la mayor terciarización del empleo, el aumento del nivel educativo de la población, el incremento de los flujos de inmigración y la descentralización de la negociación colectiva, entre otros) así como diversos avances tecnológicos que han afectado a la estructura salarial en los países desarrollados. Todo esto ha provocado un cambio en la demanda de mano de obra a favor de los trabajadores más cualificados frente a los menos cualificados (ver (Antonczyk *et al.*, 2010), (Atkinson, 2008), (Lemieux, 2008), entre otros).

En este contexto, no resulta extraño la proliferación de estudios que tratan de abordar el origen y los factores determinantes de la desigualdad salarial. Sin embargo, en los últimos años este interés se ha visto incrementado principalmente por dos aspectos relevantes y de gran actualidad. Por un lado, la reciente, duradera y profunda crisis económica (recesión) que comenzó a finales de 2007 con la quiebra de Lehman Brothers y que está teniendo grandes consecuencias tanto en los ingresos de los asalariados (sobre todo en su grado de desigualdad y en la aparición de importantes niveles de pobreza relativa asociada a bajos salarios) como en los ingresos impositivos que la Administración obtiene a través de las diferentes figuras impositivas (IRPF, Impuesto de Sociedades, Impuestos Indirectos, Cotizaciones Sociales e Ingresos no impositivos). Por otro lado, la más que previsible crisis de deuda soberana asociada al Cov-19 que habrá que hacer frente en los próximos años y su impacto tanto en la desigualdad de los individuos como en las arcas del Estado.

Con el objetivo de poder analizar la suficiencia financiera que el Sistema Fiscal español necesita para cubrir el elevado Déficit y Deuda Pública al que deberá hacer frente en los próximos años, el propio Gobierno en la Actualización del Plan de Estabilidad (Ministerio de Economía), la AIREF, la Comisión Europea y Funcas han realizado diferentes estimaciones de la elasticidad de la recaudación al ciclo económico. Sin embargo, tales elasticidades se refieren a la recaudación de todo el Sistema Fiscal y no a elasticidades concretas de las distintas figuras impositivas.

Analizando las cifras adelantadas por el Instituto Nacional de Estadística (INE) sobre el impacto que la pandemia del Cov-19 ha tenido sobre las arcas del Estado se aprecia que, a pesar del hundimiento experimentado por el PIB que alcanzó una caída del 11% en el año 2020, la recaudación fiscal mostró una notable solidez registrando una caída del 4,7% en dicho periodo. Asimismo, en el Informe Anual de Recaudación Tributaria 2020 de la AEAT se aprecia que, salvo en el caso del IRPF, cuya recaudación aumentó un 1,2% con respecto a la del año 2019, los ingresos de todas las grandes figuras tributarias disminuyeron durante la pandemia. Esto significa que el hecho de que la recaudación fiscal haya experimentado una menor volatilidad a variaciones del ciclo económico ha sido gracias al comportamiento de dos figuras tributarias: el IRPF y las cotizaciones sociales, lo que muestra el esfuerzo que han realizado los hogares para mantener los recursos públicos. Sin embargo, es de destacar que se aprecia un comportamiento claramente diferenciado en las rentas del trabajo en 2020 según fuera su origen (es decir, salarios privados, salarios públicos, pensiones, etc.), donde se puede apreciar un claro incremento de la desigualdad salarial.

En este contexto en el que los ingresos por IRPF y cotizaciones sociales representan el 45,3% del total de ingresos del Sistema Fiscal en el año 2020 y existe un aumento de la desigualdad salarial, es de gran interés para los responsables públicos conocer la evolución de la desigualdad de los ingresos de los asalariados, así como las consecuencias recaudatorias que este hecho pudiera tener sobre los ingresos fiscales, principalmente en el IRPF.

Con todo ello, el presente artículo persigue dos objetivos fundamentales. Por un lado, la estimación de las correspondientes funciones de densidad de las rentas salariales antes de IRPF que nos permitirán analizar la distribución de dichos ingresos y determinar cómo el contexto actual ha incidido en su evolución. Por otro lado, analizar cómo la desigualdad observada en las distribuciones de los ingresos salariales antes de impuestos afecta a los ingresos impositivos obtenidos por la Administración Tributaria por IRPF, dada una estructura más o menos progresiva, es decir, estimar la elasticidad de la recaudación por IRPF a variaciones en la desigualdad y en las rentas salariales.

Centrándonos en el primer objetivo de este trabajo, comenzamos analizando la evolución de las rentas salariales de los trabajadores españoles utilizando la información ofrecida por los registros administrativos de la AEAT. Con el objetivo de obtener estimadores consistentes y, al mismo tiempo, evitar problemas relacionados con la mala especificación del modelo se propone una metodología no paramétrica totalmente novedosa que permite hacer frente al hecho de que los datos utilizados aparecen recogidos en intervalos asimétricos y evitar tener que recurrir a supuestos paramétricos sobre las formas funcionales que pueden no estar avalados por los datos. Posteriormente, centrándonos en el segundo objetivo del proyecto, se establece un marco teórico no paramétrico con el objetivo de entender y poder predecir con una mayor exactitud la evolución esperada de la recaudación fiscal por IRPF ante variaciones en las rentas salariales y los niveles de desigualdad de los contribuyentes españoles. En concreto, proponemos un procedimiento de estimación no paramétrico de dos etapas. En la primera etapa, se recurre a un procedimiento de máxima verosimilitud para estimar el nivel de desigualdad (medido a través del índice de Gini), para una renta salarial dada, bajo el supuesto de que la distribución de la renta salarial es conocida y parametrizable en términos de una forma funcional (como la distribución log-normal o la distribución de Pareto) y que, a su vez, admite una transformación en términos del índice de Gini. En la

segunda etapa, se presenta un procedimiento de estimación local constante con el objetivo de estimar no paramétricamente la elasticidad fiscal por IRPF a variaciones en las rentas salariales y en los niveles de desigualdad.

El resto del trabajo se estructura de la siguiente manera. En la Sección 2 se repasa la literatura reciente sobre desigualdad salarial y sobre la relación existente entre desigualdad y recaudación por IRPF. En la Sección 3 se discuten las principales características de los datos que se van a utilizar. En la Sección 4 se presentan los distintos procedimientos de estimación no paramétricos desarrollados en este trabajo, haciendo hincapié en los principales problemas metodológicos a los que hay que hacer frente desde el punto de vista empírico. En la Sección 5 se recogen los principales resultados obtenidos aplicando la metodología desarrollada en este proyecto a los datos procedentes de registros administrativos de la AEAT. Finalmente, las conclusiones se recogen en la Sección 6.

2. REVISIÓN DE LA LITERATURA

Como se ha comentado previamente, uno de los temas de más actualidad en las últimas décadas es el aumento de la desigualdad salarial en la mayoría de los países desarrollados. En concreto, el índice de Gini de los países de la OCDE se ha visto incrementado alrededor de tres puntos porcentuales en los últimos veinticinco años (OECD, 2014). Sin embargo, este aumento de la desigualdad no se ha producido uniformemente dado que existe un comportamiento claramente diferenciado entre países, especialmente después de la crisis financiera que irrumpió a finales del año 2007.

A diferencia de lo que podía ser esperado, esta crisis financiera supuso para muchos países de la Unión Europea una estabilización de la desigualdad. Sin embargo, éste no parece ser el caso de España que ha experimentado un crecimiento continuado de la desigualdad en los últimos años, pasando de un valor del índice de Gini de alrededor de 0,319 en el año 2007 a un valor de 0,337 en el año 2013 (Eurostat, 2015), y situándose actualmente como uno de los países europeos y de la OCDE con más desigualdad en la distribución de la renta disponible. Este hecho ha sido analizado por diversos organismos internacionales como la OCDE (OECD, 2014b) y el Fondo Monetario Internacional (IMF, 2014) encontrando una relación directa entre el incremento en la desigualdad salarial y el enorme aumento experimentado por el desempleo en España desde el inicio de la crisis de 2007, así como por el comportamiento mostrado por los salarios durante dicho periodo. Sin embargo, ambas instituciones también señalan que en los últimos años se ha logrado reducir apreciablemente tanto la desigualdad de la renta disponible de los hogares como el nivel de pobreza gracias al carácter progresivo que parece haber tenido el proceso de consolidación fiscal realizado por el gobierno español a partir del año 2010. En concreto, en la última década los objetivos prioritarios del gobierno se han centrado en tratar de trasladar, en términos relativos, mayores cargas impositivas y reducciones de gasto público a los hogares con rentas más altas con el objetivo de reducir la desigualdad entre los individuos.

A pesar de la mejoría experimentada en los últimos años, actualmente España sigue adoleciendo de niveles de desigualdad de la renta elevados, situación que se prevé empeore con el contexto

actual de crisis generada por la pandemia Cov-19. En este contexto, el desarrollo de sistemas tributarios que sean capaces de proporcionar recursos suficientes para la sociedad y, al mismo tiempo, reducir la desigualdad de la renta y de la riqueza entre los individuos es totalmente esencial (Onrubia Fernández y Rodado Ruiz, 2015). Finalmente, (López-Laborda y Sanz Arcega, 2012) alegan que existe una valoración social bastante positiva respecto a la utilización del sistema fiscal como instrumento para corregir la desigualdad de la renta. Sin embargo, otros autores como (Ayala, 2013) y (Ayala *et al.*, 2013), entre otros, señalan que el aumento de la desigualdad de la renta disponible desde el inicio de la crisis de 2007 coincide con el deterioro de la capacidad redistributiva del sistema tributario español, mostrando su influencia limitada a la hora de corregir la desigualdad de la renta.

Con el objetivo de tener una panorámica global que nos permita entender la evolución de la desigualdad salarial en España y su relación con la recaudación fiscal, en esta sección se pretende realizar una revisión intensiva sobre la literatura más reciente. En primer lugar, nos centraremos únicamente en los aspectos relativos a la desigualdad salarial. Posteriormente, se extenderá el estudio a la relación existente entre desigualdad y recaudación fiscal.

2.1. Literatura sobre la desigualdad salarial

El origen de la desigualdad salarial ha sido intensamente estudiado en las tres últimas décadas, especialmente en el contexto internacional, como se aprecia en las revisiones de (Blau y Kahn, 1999) y (Katz y Autor, 1999), por ejemplo. Sin embargo, el análisis de los factores determinantes de la desigualdad salarial en España ha sido bastante escaso, aunque en la última década ha aumentado el interés por tales cuestiones.

La estimación de las funciones de distribución de los ingresos (o curvas de Lorenz) depende esencialmente de los datos disponibles, y ésta es una de las razones principales para la escasez de estudios sobre la economía española. Inicialmente, la mayor parte de estos estudios utilizaban la información recogida en la antigua Encuesta de Presupuestos Familiares (EPF), en las sucesivas oleadas de la Encuesta Continua de Presupuestos Familiares (ECPF) o en el European Community Household Panel (ECHP). Posteriormente, muchos de ellos recurrieron a la Encuesta de Estructura Salarial (EES) del Instituto Nacional de Estadística (INE) y a la Muestra Continua de Vidas Laborales (MCVL) de la Seguridad Social.

La mayoría de estas encuestas se caracterizan por una amplia discrecionalidad temporal en sus publicaciones, así como por la dificultad de deducir los ingresos salariales de los miembros del hogar con la información que proporcionan. Para resolver esta situación, otras fuentes de información procedentes de los registros administrativos como los proporcionados por la Agencia Tributaria a través de la publicación “Empleo, Salarios y Pensiones en las Fuentes Tributarias” (ESP) o la MCVL generada a partir de los registros administrativos de la Seguridad Social parecen ser fuentes de información más recomendables dado que ofrecen información salarial con periodicidad anual. Sin embargo, la utilización de la información recogida por estos registros administrativos no es una tarea sencilla dado que estos registros suministran los datos agrupados en intervalos que pueden

ser simétricos (cuando el número de individuos en cada intervalo es el mismo) o asimétricos (cuando el número de individuos por intervalo es distinto).

La estimación de las funciones de densidad de frecuencia y/o curvas de Lorenz de los salarios utilizando datos agrupados puede realizarse a través de dos enfoques: paramétrico o no paramétrico. El enfoque paramétrico consiste en construir la curva de Lorenz a partir de funciones de densidad continuas a la distribución de la renta de las cuales se puedan derivar los indicadores de desigualdad o de pobreza (Walter y Weimer, 2018). Las primeras funciones que se propusieron como candidatas para describir la distribución de los ingresos, típicamente asimétrica hacia la derecha, fueron la distribución de Pareto y la distribución log-normal. Sin embargo, los ajustes que proporcionan no son muy buenos, prefiriéndose la función de Pareto para caracterizar las colas altas de la distribución, y la log-normal para los tramos medios. Posteriormente, se han propuesto una gran diversidad de funciones de distribución cuya complejidad ha ido creciendo progresivamente (distribución de Fisk, Weibull, gamma, gamma generalizada, etc.) pero (Kleiber, 2008) o (Reed y Wu, 2008), entre otros, argumentan que las curvas de Lorenz basadas en la función de verosimilitud de una distribución multinomial proporcionan mejores resultados cuando los datos están agrupados en intervalos (ver (Bandourian, *et al.*, 2003), (Kleiber y Kotz, 2003) y (Reed y Wu, 2008) para una mayor discusión a este respecto). Sin embargo, este enfoque paramétrico ha recibido diversas críticas, como se recoge en (Kakwani y Podder, 2008), dado que el comportamiento de estimador es bastante pobre si el modelo presenta un error de especificación (es decir, si la función de densidad continua que se especifica no está avalada por los datos).

Como solución a este tipo de problemas surgen las técnicas de estimación no paramétricas de la densidad que se caracterizan por estimar la función de distribución a través del uso de kernels, es decir, sin presuponer ninguna forma funcional a priori. De este modo, estas estimaciones son robustas a posibles errores de especificación de la función a estimar (ver (Silverman, 1986) para profundizar en estas técnicas). En los últimos años este tipo de técnicas no paramétricas han sido utilizadas en distintos ámbitos como el estudio de la pobreza y la desigualdad (ver (Sala-i-Martín, 2002), (Sala-i-Martín, 2006), (Fuentes, 2005), entre otros). Es de destacar que la precisión de estos resultados depende esencialmente de las funciones de kernels y del ancho de ventana (o bandwidth) utilizados en el cálculo de la función de densidad, especialmente cuando los datos agrupados son la fuente de información (ver (Minou y Reedy, 2014), (Wu y Perloff, 2003)). Asimismo, está ampliamente demostrado que estos procedimientos no paramétricos funcionan considerablemente bien cuando las observaciones disponibles son elevadas. Desgraciadamente, en el caso de los salarios este requisito no se cumple dado que, a pesar de ser una variable continua, está agrupada en intervalos. Una posible solución a este problema es la utilización de procedimientos semiparamétricos que combinan la especificación paramétrica con una parte no paramétrica, manteniendo de este modo la flexibilidad que ofrecen este último tipo de modelos. En este caso modelos de dos etapas como los propuestos en (Moral-Arce y De las Heras, 2011) y (Moral-Arce, *et al.*, 2019) son más adecuados. En concreto, estos autores proponen transformar la información proporcionada en cuantiles (insuficiente información para la estimación de la función de densidad no paramétrica a partir de técnicas estándar) en un vector que contiene todos los valores ficticios de ingreso necesarios para estimar la función de densidad de una manera más adecuada. En la

primera etapa, se estima una curva de Lorenz usando la información agrupada en intervalos y se predicen “N” valores de ingreso de los individuos. Con estos valores, en la segunda etapa se estima la distribución del ingreso a través del uso de los estimadores de la densidad de kernel no paramétricos.

En la Tabla 1 se resumen algunos de los estudios más relevantes sobre los factores determinantes de los niveles de desigualdad salarial en España, destacando el periodo analizado, la fuente de datos utilizada y el procedimiento de estimación empleado.

Tabla 1
RESUMEN PRINCIPALES TRABAJOS DESIGUALDAD SALARIAL EN ESPAÑA

Artículo	Base de datos	Periodo	Enfoque
(Arellano, et al., 2002)	EPA	1987-1994	Paramétrico
(Bonhomme y Hospido, 2012a)	ESP	2004-2010	No paramétrico
(Bonhomme y Hospido, 2012b)	MCVL	1888-2010	No paramétrico
(Bonhomme y Hospido, 2017)	MCVL	1988-2010	Paramétrico
(Carrasco, et al., 2011)	EES	1995-2002	Paramétrico
(Hidalgo, 2008)	EPF	1980-2000	Paramétrico
(Hidalgo, 2010)	EPF	1980-2000	Paramétrico
(Izquierdo y Lacuesta, 2007)	EES	1995-2002	Paramétrico
(Izquierdo y Lacuesta, 2012)	EES	1995-2002	Paramétrico
(Moral-Arce y De las Heras, 2011)	ESP	2003	Semiparamétrico
(Palacio y Simón, 2004)	EES	1995	Paramétrico
(Pijoan y Sánchez-Marcos, 2010)	ECPF, ECHP	1992-2001	Paramétrico
(Quilis y Frutos, 2020)	ESP	2001-2017	Paramétrico
(Simón, 2009)	EES	1995-2002	Paramétrico

Fuente: elaboración propia.

Analizando en detalle los trabajos más recientes recogidos en la Tabla 1, y a pesar de las diferentes fuentes de información utilizadas y de los distintos periodos considerados, es posible esbozar una serie de rasgos comunes sobre la evolución y las principales causas de la desigualdad salarial en España. Por un lado, la desigualdad de los ingresos salariales y de la renta en España presentan un marcado componente contracíclico dado que experimentan una fuerte caída en la parte baja de la distribución a lo largo de los períodos recesivos y aumentan en las épocas de bonanza económica (ver (Bonhomme y Hospido, 2017), (Carrasco et al., 2015), (Izquierdo y Lacuesta, 2012) o (Pijoan y Sánchez-Marcos, 2010), entre otros). Por otro lado, los cambios en la composición de la población activa y, sobre todo, el papel de las empresas ayudaron a disminuir considerablemente la desigualdad salarial durante el periodo 1995-2006 (ver (Izquierdo y Lacuesta, 2012), (Carrasco et al., 2011) y (Simón, 2009), entre otros). Finalmente, la demanda de trabajadores altamente cualificados también muestra un comportamiento contracíclico. Durante el periodo 1985-1992, donde se produjo una reducción del desempleo, la desigualdad salarial cayó como consecuencia de la reducción del “premio” salarial que recibían los trabajadores altamente cualificados (Pijoan

y Sánchez-Marcos, 2010). Por el contrario, desde el inicio de la crisis de 2007, donde se produjo un gran aumento del desempleo, la desigualdad salarial se ha visto incrementada notablemente como consecuencia de la mayor demanda de trabajadores altamente cualificados (Carrasco *et al.*, 2015). Una discusión mucho más detallada sobre estos aspectos se encuentra en (García-Serrano y Arranz, 2013).

Asimismo, como se aprecia en la Tabla 1, la mayoría de los estudios que analizan el caso español se centran en enfoques paramétricos. En los últimos años existe un interés creciente en la extensión de estos estudios al caso no paramétrico como se aprecia en (Bonhomme y Hospido, 2012a), (Bonhomme y Hospido, 2012b) y al semiparamétrico (Moral-Arce y De las Heras, 2011). Sin embargo, a pesar de los esfuerzos realizados en los últimos años, todavía existe mucho trabajo por realizar en lo que respecta a la estimación no paramétrica de la densidad cuando los datos están agrupados. Para resolver esta situación, en este artículo se desarrolla una técnica de estimación no paramétrica para datos agrupados que proporciona estimaciones de las curvas de Lorenz con buenas propiedades estadísticas en una sola etapa.

2.2. Literatura sobre desigualdad e ingresos por IRPF

En los últimos años buena parte de los países avanzados se encuentran implementando políticas fiscales redistributivas según las cuales los ingresos personales (salarios) se gravan a una tasa cada vez mayor, mientras que las transferencias tienden a dirigirse a los hogares con menores recursos. En este contexto, existe un intenso debate en España sobre cómo financiar los planes de recuperación del estímulo fiscal para paliar las consecuencias económicas de la crisis Cov-19 y, más precisamente, sobre cómo afrontar el inevitable y necesario proceso de consolidación fiscal que muy probablemente tendrá que llevar cabo el gobierno español.

Actualmente se puede identificar un movimiento social claramente diferenciado que solicita la implantación de la llamada “justicia fiscal” que consiste en solicitar un aumento de la progresividad del IRPF a los contribuyentes con mayores salarios con el objetivo de conseguir un aumento de la recaudación fiscal. Sin embargo, este tipo de políticas podría resultar muy costoso en términos de eficiencia, tal y como se destaca en (Serrano-Puente, 2020), dado que el principal objetivo de los responsables públicos debe ser que el sistema tributario sea proporcional, es decir, que las desigualdades en el ingreso antes y después de impuestos fuesen idénticas. Por todo ello, es totalmente necesario poder determinar si la desigualdad en la distribución de los salarios realmente afecta a la recaudación por IRPF, de qué forma y en qué cuantía.

En este contexto, la literatura económica reciente que relaciona la distribución de las rentas salariales y los ingresos fiscales se ha centrado principalmente en dar respuesta a tres preguntas esenciales: (i) ¿Cómo la progresividad del Sistema Fiscal afecta a la distribución de los ingresos (salarios) después del impuesto y, por lo tanto, disminuye la desigualdad entre la renta antes y después del impuesto? (ii) ¿Cómo los cambios en el sistema impositivo (a través de diversas reformas fiscales) afectan a los ingresos impositivos y a la capacidad redistributiva del mismo? (iii) ¿Cómo la desigualdad observada en la distribución de los ingresos salariales afecta a los ingresos públicos?

Con el objetivo de dar respuesta a la primera cuestión, así como analizar las ganancias en bienestar social asociadas a la progresividad del sistema, se suele recurrir a la utilización de muestras de contribuyentes de gran tamaño a lo largo del tiempo (microdatos). Este tipo de información ha permitido construir una amplia base de datos de panel sobre los contribuyentes (ver (Onrubia y Picos, 2011) y (Onrubia y Picos, 2012), para más información sobre sus características en el caso español) que posibilita a los investigadores el desarrollo de modelos de microsimulación que permiten analizar los efectos de las distintas reformas fiscales sobre el IRPF (ver (De las Heras, et al., 2004), (Díaz de Sarralde et al., 2006), (Mercader-Prats et al., 2001), (Onrubia y Sanz, 2003), (Oliver y Spadaro, 2009), (Picos et al., 2009) y (Sanz et al., 2008), por citar solo algunos casos para el caso español). En concreto, estos modelos de microsimulación (con y sin comportamiento) persiguen dos objetivos fundamentales. Por un lado, tratan de determinar si los cambios en la estructura impositiva implican o no cambios en el comportamiento económico de los contribuyentes a partir del concepto de elasticidad de la oferta de trabajo (ver (Blundell et al., 1988), (Blundell et al., 1998), (Bourguignon y Magnac, 1990), (Hausman, 1980), (Hausman y Ruud, 1984), entre otros). De este modo, estos artículos demuestran cómo cambios en la estructura fiscal puede inducir a los individuos a alterar las horas de trabajo ofertadas modificando sus ingresos antes de impuestos, así como la recaudación fiscal correspondiente. Por otro lado, analizar el efecto recaudatorio y redistributivo de dichos cambios en las estructuras fiscales, tanto directos como indirectos (Bourguignon y Spadaro, 2005).

Centrándonos en la segunda cuestión, la literatura más reciente sobre las respuestas de los contribuyentes ante cambios en la legislación fiscal ha pasado de un enfoque basado en las elasticidades de la oferta de trabajo hacia otro basado en las elasticidades de la renta gravable a los tipos marginales. Este enfoque alternativo, también conocido como “Elasticity of Taxable Income” (ETI), fue propuesto inicialmente por (Feldstein, 1995) y (Feldstein, 1999) como una medida del conjunto de respuestas de los contribuyentes (entre ellos los asalariados) ante cambios en los tipos marginales y proporciona un indicador estadístico para el análisis de la recaudación y de los efectos sobre la eficiencia del impuesto sobre la renta.

A este respecto cabe destacar que el ETI es una medida muy atractiva desde el punto de vista empírico dado que permite cuantificar todas las posibles respuestas que la elasticidad de la oferta de trabajo no recogía. En concreto, la renta gravable puede variar por diversos motivos: (i) como respuesta a los cambios en los tipos marginales del impuesto si los contribuyentes deciden trabajar más, mejor o más intensamente para obtener una mayor renta, (ii) porque declaren una mayor proporción de sus ingresos (evasión fiscal), (iii) porque pueden efectuar un intercambio entre sus rentas del trabajo y del capital para minimizar la presión fiscal. Además, el ETI recoge aspectos tan diversos como el nivel de ahorro, la forma en que se perciben las rentas, la elección entre tipos de activos, la elección entre inversión y consumo, el fraude fiscal, y los posibles ajustes intertemporales en la renta que puedan efectuar los contribuyentes, por lo que es un parámetro clave en temas tales como la redistribución óptima de la renta.

En los últimos años el ETI ha experimentado un importante desarrollo, tanto desde el punto de vista teórico como empírico, con el objetivo de analizar las respuestas de los contribuyentes a los

cambios en los tipos marginales (ver (Sáez, 2001), (Gruber y Sáez, 2002), (Chetty, 2009), (Creedy, 2009), (Creedy, 2010), (Giertz, 2009), (Sáez *et al.*, 2012), (Creedy y Gemmel, 2013), por ejemplo).

Considerando la tercera cuestión, la literatura económica existente se ha centrado en analizar cómo deberían ser los sistemas de imposición óptima en un marco de crecientes desigualdades (ver (Poterba, 2007), (Slemrod y Bakija, 2001), (Gan y Tan, 2011), entre otros) o cómo la reciente desigualdad puede afectar a los programas de gasto público de los distintos gobiernos (Boustan, *et al.*, 2013). En concreto, estos trabajos señalan el incremento de la desigualdad en los ingresos de los individuos desde la década de los ochenta y por tanto su importancia en la variación de las bases imponibles en todos los países. Sin embargo, el análisis cuantitativo sobre los efectos de la creciente desigualdad en los ingresos impositivos y/o su relación no ha sido objeto generalizado de investigación. Recientemente, (De las Heras y Rodríguez-Poo, 2015) desarrollan un procedimiento de estimación que permite establecer cierto tipo de relación funcional entre la desigualdad de los ingresos salariales y la recaudación fiscal por IRPF para el periodo 2006-2012. No obstante, se desconoce cómo el impacto de la crisis económica ha podido afectar a dichas relaciones en los últimos años y si el procedimiento de estimación propuesto por dichos autores es directamente extrapolable al marco actual.

Finalmente, es necesario destacar un aspecto muy relevante desde el punto de vista metodológico. Tanto en los modelos de microsimulación con comportamiento como en las estimaciones de las elasticidades, ya sean sobre la oferta de trabajo como sobre la elasticidad de la renta gravable, se considera la oferta de trabajo (y por tanto la desigualdad salarial) como una variable endógena a los propios modelos y a los parámetros a estimar. En otras palabras, se supone que los cambios en la oferta de trabajo en estos modelos dependen de los cambios en la estructura de los sistemas impositivos. Sin embargo, tal y como se indica en (Poterba, 2007), “entender la relación entre la política fiscal y la generación de ingresos es importante para entender la dinámica de la distribución del ingreso. Además, entender la evolución de la distribución de ingresos antes de impuestos también es esencial para hacer un pronóstico de ingresos preciso”.

En este contexto, la segunda parte de este artículo se centra en el desarrollo de una técnica de estimación más flexible que permita analizar la relación existente entre la desigualdad salarial y los ingresos por IRPF y corroborar si realmente se aprecia un comportamiento diferente al destacado para el periodo 2006-2012 utilizando datos agrupados procedentes de registros administrativos públicos.

3. DATOS

La decisión sobre cuál es el método de estimación más apropiado depende crucialmente de cómo han sido agrupadas las observaciones y de cómo la información relevante ha sido puesta a disposición de los investigadores. Por lo tanto, en esta sección se realizará una pequeña introducción de las distintas fuentes de información que se van a utilizar en este artículo.

Los estudios sobre desigualdad salarial suelen recurrir a encuestas estadísticas que tratan de caracterizar el comportamiento de la población a partir de la información proporcionada por un

conjunto de hogares seleccionados por muestreo aleatorio, como la Encuesta de Condiciones de Vida o la Encuesta de Población Activa elaboradas por el INE. Sin embargo, esta caracterización de la población se realiza a través de una serie de estimaciones que dependen de supuestos claves sobre la operación de muestreo, funciones de distribución, etc., que pueden ser objeto de debate.

Para evitar este tipo de situaciones a la hora de analizar la desigualdad de la renta en España, en este artículo se utilizan dos fuentes estadísticas: “Mercado de Trabajo y Pensiones en las Fuentes Tributarias”, publicada por la AEAT desde el año 1999 hasta la actualidad¹, y “Estadística de los Declarantes del IRPF”, publicada por la AEAT desde el año 2003 hasta la actualidad².

Por un lado, para analizar la distribución de los salarios en el mercado español a lo largo del tiempo utilizaremos la información recogida en la publicación “Mercado de Trabajo y Pensiones en las Fuentes Tributarias”. Esta estadística tiene como principal fuente de información las declaraciones anuales de los principales perceptores de renta de modo que resume las retenciones e ingresos a cuenta practicados por las entidades pagadoras de salarios, pensiones, prestaciones por desempleo y otras rentas³ que tienen domicilio fiscal en el Territorio de Régimen Fiscal Común, con independencia de cuál sea la Administración (estatal o foral) en la que declare su empleador. Por lo tanto, la información recogida por esta publicación es de gran utilidad dado que nos permite conocer la distribución demográfica y territorial de la renta del trabajo, así como otras características del mercado de trabajo. En nuestro caso concreto nos centraremos en la información recogida para el territorio de régimen común durante el periodo 1999-2020.

En concreto, la Tabla 2 muestra la información disponible sobre los asalariados (desde la óptica del perceptor) para el año 1999, pero esta estructura se mantiene para el resto de años.

Como se aprecia en la Tabla 2, la información disponible sobre asalariados y rendimientos salariales se encuentra agrupada en intervalos ordenados de manera ascendente según ciertos niveles de rendimiento que dependen del Salario Mínimo Interprofesional (SMI). Asimismo, como suele ser habitual cuando la información de origen procede de recursos administrativos, los intervalos de ingresos son asimétricos (es decir, el número de individuos en cada intervalo es diferente) por lo que no es posible estimar directamente la función de densidad de la renta basada en cuantiles como en (Sala-i-Martín, 2006), entre otros.

¹ La publicación completa desde 1999 se puede consultar en:

https://www.agenciatributaria.es/AEAT.internet/datosabiertos/catalogo/hacienda/Mercado_de_Trabajo_y_Pensiones_en_las_Fuentes_Tributarias.shtml

² La publicación completa desde 2003 se puede consultar en:

https://sede.agenciatributaria.gob.es/Sede/datosabiertos/catalogo/hacienda/Estadistica_de_los_declarantes_del_IRPF.shtml?faqlId=a119ec15d48c0710VgnVCM100000dc381e0aRCRD

³ La estadística “Mercado de Trabajo y Pensiones en las Fuentes Tributarias” contiene información sobre los asalariados, desempleados que perciben prestaciones por desempleo y pensionistas, personas que han abandonado el mercado laboral (por razón de edad o incapacidad) o que perciben pensiones de viudedad.

Tabla 2

**VALORES Y DISTRIBUCIÓN DE ASALARIADOS Y SALARIOS TOTALES. TERRITORIOS DE RÉGIMEN COMÚN.
AÑO 1999**

Tramos de salario	Asalariados	Distribución porcentual asalariados	Salarios	Distribución porcentual de salarios
De 0 a 0,5 SMI	2.783.810	19,29	3.427.588.455	1,87
De 0,5 a 1 SMI	1.688.228	11,70	7.362.534.947	4,01
De 1 a 1,5 SMI	1.743.305	12,08	12.778.091.149	6,96
De 1,5 a 2 SMI	2.182.373	15,12	22.269.397.014	12,13
De 2 a 2,5 SMI	1.595.219	11,05	20.730.962.502	11,29
De 2,5 a 3 SMI	1.033.223	7,16	16.479.668.049	8,98
De 3 a 3,5 SMI	799.476	5,54	15.096.896.660	8,22
De 3,5 a 4 SMI	642.458	4,45	14.006.579.842	7,63
De 4 a 4,5 SMI	523.604	3,63	12.931.235.840	7,04
De 4,5 a 5 SMI	373.810	2,59	10.305.239.854	5,61
De 5 a 7,5 SMI	718.098	4,98	24.947.558.001	13,59
De 7,5 a 10 SMI	204.777	1,42	10.172.993.259	5,54
Más de 10 SMI	142.888	0,99	13.048.716.371	7,11
Total	14.431.269	100	183.557.461.943	100

Fuente: Estadística "Mercado de trabajo y pensiones en las fuentes tributarias". AEAT. SMI: Salario Mínimo Interprofesional.

Por su parte, la Tabla 3 recoge una representación general de la agrupación de datos procedentes de registros administrativos que es necesario tener en cuenta a la hora de desarrollar un procedimiento de estimación que nos permita estimar consistentemente la distribución de la renta de los hogares españoles.

Tabla 3

DATOS RELATIVOS Y ACUMULADOS SOBRE INGRESOS Y ASALARIADOS AGRUPADOS

Intervalos	0-x ₁	x ₁ -x ₂	...	x _{j-1} -x _j	...	x _{j-1}	Soporte total
Asalariados	n ₁	n ₂	...	n _j	...	n _j	N
Salarios	z ₁	z ₂	...	z _j	...	z _j	Z
Asalariados acum.	$p_1 = \frac{n_1}{N}$	$p_2 = \frac{n_1 + n_2}{N}$...	$p_j = \frac{n_1 + \dots + n_j}{N}$...	$p_j = 1$	100%

Fuente: elaboración propia.

Por otro lado, con el objetivo de obtener información sobre los ingresos impositivos por IRPF recurrimos a los datos recogidos en la publicación "Estadística de los Declarantes del IRPF". Esta estadística está basada en las declaraciones del IRPF, correspondiente al ejercicio fiscal de referencia de la estadística, que presentan los contribuyentes con domicilio fiscal en el Territorio de Régimen Fiscal Común. De este modo, recoge la información detallada de las principales partidas

contenidas en el modelo de declaración D-100 de todos aquellos perceptores de rentas que mantienen una relación fiscal con la Administración y permite la realización de diversos cruces de información entre perceptores-declarantes y pagadores de rentas que no sería posible con microdatos procedentes de estadísticas oficiales.

Al igual que en la publicación anterior, la información se ofrece en intervalos de renta cuya estructura asimétrica debe ser tenida en cuenta para obtener resultados consistentes. Asimismo, cabe destacar que la información recogida por esta estadística permite establecer una relación entre la variable de renta salarial y el Total de Ingresos Computables o entre la variable de Ingresos fiscales por IRPF (epígrafe 742 de la estadística denominada “Retención por rendimientos del trabajo”) y la Retención y/o Cuota Líquida por rendimientos del trabajo. En ambos casos las variables fiscales se presentan agrupadas por tramos de rendimiento y, para cada tramo, conocemos el número de declarantes, valor medio de dicha variable y frecuencias absolutas de declarantes. De este modo, la información proporcionada por esta estadística es de gran utilidad a la hora de establecer la elasticidad de la recaudación fiscal a cambios en la distribución de la renta y/o la retribución salarial.

4. METODOLOGÍA Y PROCEDIMIENTO DE ESTIMACIÓN

Tal y como se ha comentado previamente, en este artículo se persiguen dos objetivos claramente diferenciados desde el punto de vista metodológico. Por un lado, el desarrollo de un procedimiento de estimación no paramétrico que permita estimar directamente las funciones de densidad con datos agrupados en intervalos asimétricos. Por otro lado, proponer un procedimiento semiparamétrico y/o no paramétrico que permita establecer algún nexo funcional entre la desigualdad salarial y la recaudación impositiva por IRPF. En las siguientes subsecciones se desarrollará cada uno de estos procedimientos por separado.

4.1. Estimación no paramétrica de la densidad con datos agrupados

En esta subsección se presentará una técnica de estimación no paramétrica para estimar una función de densidad que nos permita analizar la desigualdad de las rentas salariales y de la recaudación fiscal bajo dos escenarios alternativos:

- a) Continuidad de la renta y de los ingresos fiscales.
- b) Observabilidad parcial de dichas variables en intervalos.

Centrándonos en el primer escenario denotamos por Y_1, Y_2, \dots, Y_N las rentas salariales de N individuos y suponemos que dichas rentas están caracterizadas por una función de densidad de probabilidad $f(y)$ totalmente desconocida. Para estimar esta función y sus funcionales, tales como la curva de Lorenz y el índice de Gini, podemos utilizar el siguiente estimador de la densidad de kernel que es una de las técnicas de estimación no paramétricas más conocidas en la literatura y que fue introducida inicialmente por (Rosenblatt, 1956) y (Parzen, 1962),

$$\check{f}_h(y) = \frac{1}{Nh} \sum_{i=1}^N K\left(\frac{y-Y_i}{h}\right), \quad (1)$$

donde (y) es el punto de evaluación, $h > 0$ es conocido como el ancho de ventana o parámetro de suavizado y $K(\cdot)$ es la función de núcleo o kernel comúnmente utilizada en las técnicas de estimación no paramétricas de la densidad y de la regresión.

Como se ha comentado previamente, la forma y el comportamiento de este estimador depende crucialmente de la elección del ancho de ventana y de la función de kernel⁴. En la literatura se encuentran distintas técnicas para la estimación del ancho de ventana, donde el enfoque plug-in y la validación cruzada suelen ser los predominantes (ver (Henderson y Parmeter, 2015), (Jones, et al., 1996) o (Wand y Jones, 1995), entre otros) para una revisión más intensiva al respecto. Sin embargo, la elección de la función de kernel es más compleja dado que los procedimientos de estimación no paramétricos tradicionales están sujetos a una serie de requerimientos. En concreto, las funciones de kernel más utilizadas en esta literatura (Normal, Epanechnikov,...) suelen ser simétricas, de modo que su aplicación en los datos salariales puede ocasionarnos un serio problema de sesgo en el límite (i.e. $y=0$) dado que la distribución $f(y)$ tiene soporte sólo sobre valores positivos y la utilización de kernels simétricos ponderará positivamente observaciones que quedan fuera del soporte de la densidad cuando estamos suavizando cerca del cero (ver (Silverman, 1986) para una discusión más detallada). Para resolver esta situación es necesario utilizar kernels que nunca asignen pesos fuera del soporte como la función Gamma kernel introducida en (Chen, 2000). De este modo, el estimador de la densidad resultante es de la forma

$$\check{f}_h(y) = N^{-1} \sum_{i=1}^N K_{y/h+1,h}(Y_i), \quad (2)$$

donde $K_{y/h+1,h}(Y_i)$ es la función de densidad de una variable aleatoria Gamma⁵ y h es el parámetro de suavizado que cumple las condiciones estándar (i.e., $h \rightarrow 0$ y $Nh \rightarrow \infty$ cuando $N \rightarrow \infty$). (Chen, 2000) demuestra que los estimadores de kernel gamma de la forma (2) resuelven el problema de sesgo en el límite, son no negativos y alcanzan la óptima tasa de convergencia de este tipo de problemas. De este modo, el estimador de la densidad (2) es un buen candidato para calcular la curva de Lorenz y otros indicadores de desigualdad cuando existe continuidad en la variable de salarios o de recaudación fiscal.

⁴ De manera general, la función de kernel puede ser cualquier función de densidad que cumpla las siguientes propiedades:

- $\int K(u)du = 1.$
- $\int uK(u)du = 0.$
- $\int u^2K(u)du < \infty.$

⁵ $K_{y/h+1,h}(Y_i)$ es la función de densidad de una variable aleatoria Gamma con $(y/h + 1)$ y (h) grados de libertad de la forma

$$K_{y/h+1,h}(u) = \frac{u^{y/h} e^{-u/h}}{h^{y/h+1} \Gamma(y/h + 1)}$$

y donde $\Gamma(\cdot)$ es la función Gamma estándar de la forma $\Gamma(y/h + 1) = \int_0^\infty u^{y/h} e^{-y} dy.$

Sin embargo, como se ha mostrado en la Sección 3, la información sobre los salarios y la recaudación fiscal suele estar disponible para los investigadores a través de intervalos, por lo que un procedimiento de estimación alternativo es necesario. El estudio de la estimación de la densidad con datos agrupados ha sido objeto de una intensa investigación y diversos procedimientos han sido propuestos. Por un lado, (Wang y Wertenleki, 2013) proponen un estimador de kernel agrupado no lineal a través de un procedimiento bootstrap, mientras que (Blower y Kelsall, 2002) desarrollan un estimador de kernel agrupado no lineal. Por otro lado, otro tipo de estudios proponen convertir el problema de la estimación de la densidad en un problema de regresión usando el algoritmo “root-unroot” (Brown *et al.*, 2010) o usando métodos paramétricos (Wang y Wertenleki, 2013). Recientemente, (Reyes *et al.*, 2016) y (Reyes *et al.*, 2019) proponen una modificación simple del estimador de la densidad (1).

Suponiendo que la información sobre la variable de interés (Y) se proporciona a través de un conjunto de intervalos (A_{j-1}, A_j) , donde $j=1,2,\dots,J$ representa el número de intervalos, mientras que el valor real de (Y) a nivel individual es desconocido. Definimos la longitud del intervalo j -th como $l_j = A_j - A_{j-1}$, su punto medio como $\bar{l}_j = 0,5(A_{j-1} + A_j)$, el número de observaciones dentro de cada intervalo como (N_1, N_2, \dots, N_J) , y sus correspondientes proporciones muestrales como (w_1, w_2, \dots, w_N) , donde $w_j = N_j/N$. De este modo, si el intervalo es constante el estimador (1) se puede redefinir como proponen (Scott y Sheather, 1985) y (Titterton, 1983) obteniendo la siguiente expresión

$$\tilde{f}_h^a(y) = \frac{1}{Nh} \sum_{j=1}^J N_j K\left(\frac{y-\bar{l}_j}{h}\right), \quad (3)$$

mientras que (Reyes *et al.*, 2016) extienden este estimador al caso general en el que la longitud de los intervalos es distinta obteniendo

$$\tilde{f}_h^b(y) = \frac{1}{h} \sum_{j=1}^J p_j K\left(\frac{y-\bar{l}_j}{h}\right), \quad (4)$$

donde (p_j) es el porcentaje de población acumulada como se muestra en la Tabla 3.

Sin embargo, a pesar de que (Reyes *et al.*, 2016) y (Reyes *et al.*, 2019) demuestran que el estimador de la densidad (4) presenta buenas propiedades asintóticas y en muestras finitas, es necesario seguir teniendo en cuenta que la variable de interés en nuestro caso (es decir, los salarios) solamente tiene soporte positivo y el estimador (4) está sujeto a los mismos problemas en el límite que los comentados previamente para el estimador (1). Para resolver esta situación, en este proyecto proponemos utilizar funciones asimétricas para $K(\cdot)$ como la función Gamma de modo que el estimador final que proponemos utilizar tiene la siguiente forma

$$\hat{f}_h(y) = \sum_{j=1}^J p_j K_{\bar{l}_j/h+1,h}(Y_j), \quad (5)$$

donde $K_{\bar{l}_j/h+1,h}(y) = \frac{y^{\bar{l}_j/h} e^{-y/h}}{h^{\bar{l}_j/h+1} \Gamma(\bar{l}_j/h+1)}$.

Revisando la literatura no se encuentra estudio alguno que considere las propiedades asintóticas del estimador (5). Sin embargo, combinando las estructuras de prueba de (Chen, 2000) y (Reyes *et al.*, 2019) es posible demostrar que el estimador de la densidad de kernel (5) tiene buenas propiedades asintóticas (es decir, este estimador no exhibe sesgo en el límite y alcanza la tasa óptima de convergencia para el error cuadrático medio integrado). El estudio concreto de estas propiedades está fuera del objetivo de este proyecto y se deja pendiente para una investigación futura.

Finalmente, en lo que concierne al procedimiento de selección del ancho de ventana para este caso con datos agrupados hemos de destacar que (Reyes *et al.*, 2017) desarrollan dos procedimientos alternativos (un procedimiento de plug-in y otro basado en técnicas de bootstrap) y destacan que el ancho de ventana puede ser aproximado por la siguiente expresión: $h = 0.78\sigma J^{-2/13}$, donde σ es la desviación típica de la variable \bar{I}_j y J es el número total de intervalos.

4.2. Desigualdad salarial y recaudación fiscal por IRPF

En esta sección pretendemos analizar la elasticidad de la recaudación fiscal por IRPF al ciclo económico con el objetivo de poder determinar la suficiencia financiera que el Sistema Fiscal español necesita para hacer frente a los servicios esenciales del estado del bienestar. En primer lugar, nos centraremos en el desarrollo de un estimador no paramétrico que permita calcular la elasticidad de los rendimientos fiscales a cambios en la distribución de la renta. En segundo lugar, propondremos un procedimiento de estimación flexible que permita establecer la relación entre la recaudación fiscal por IRPF y la desigualdad salarial.

4.2.1. Elasticidad fiscal y distribución de la renta

Tal y como se destaca en (Creedy, 2010), la forma de la distribución de las rentas salariales juega un papel crucial a la hora de obtener expresiones analíticas simples sobre la recaudación fiscal por IRPF. De este modo, una vez efectuada una propuesta de estimación no paramétrica de la distribución de la renta cuando la información se presenta en datos agrupados, en esta sección se establece un marco teórico que permite establecer una relación funcional entre los ingresos por IRPF y los rendimientos salariales y que nos permitirá estimar la elasticidad de la recaudación fiscal con respecto a variaciones en los salarios y en la desigualdad salarial. Posteriormente, nos centraremos en detallar el procedimiento de estimación que nos permite calcular dichas elasticidades.

Con el objetivo de calcular los rendimientos fiscales por IRPF suele ser habitual asumir que los ingresos fiscales totales presentan formas funcionales lineales como por ejemplo la “función de impuestos de varias etapas” (ver (Creedy y Gemmell, 2006)). Sin embargo, como se ha comentado previamente, estas formas funcionales especificadas a priori sobre la estructura fiscal pueden conducirnos a un problema de especificación de la forma funcional cuando ésta no está avalada por los datos que nos puede conducir a la obtención de estimadores inconsistentes. Para evitar este tipo de problemas, en este artículo proponemos una especificación más flexible de la función de los rendimientos fiscales, $R(\cdot)$, permitiendo que sean los propios datos los que determinen la forma

funcional de la distribución de la recaudación y sin tener que suponer que dicha función pertenece a un espacio de funciones paramétricas.

Siendo Y la renta salarial de los contribuyentes, T los ingresos fiscales por IRPF (medidos a través de la cuota líquida y/o las retenciones), y suponiendo que la variable Y sigue una función de densidad de probabilidad $f(y)$. Con el objetivo de calcular los rendimientos fiscales por IRPF proponemos el siguiente modelo no paramétrico

$$T=R(Y)+e \quad (6)$$

donde (e) representa el error idiosincrático y $R(\cdot)$ es una función totalmente desconocida. Tomando la esperanza condicionada en ambos lados de (6) con respecto a un valor de renta dado ($Y=y$) y asumiendo que $E(e|Y=y)=0$, podemos expresar la función de ingresos fiscales por IRPF del siguiente modo

$$R(y) = E(T|Y = y) = \int_0^{\infty} \tau f\left(\frac{\tau}{y}\right) d\tau, \quad (7)$$

donde $R(y)$ representa el ingreso fiscal por IRPF asociado al ingreso gravable, es decir, los impuestos pagados cuando el nivel de renta está por encima de un umbral dado ($y>a$), τ representa los ingresos fiscales por IRPF y $f\left(\frac{\tau}{y}\right)$ es la función de densidad de probabilidad de los ingresos fiscales por IRPF condicionada al nivel de renta ($Y=y$). Por su parte, el total de la recaudación fiscal por IRPF (R) se representa como

$$R = \int_0^{\infty} R(y)f(y)dy. \quad (8)$$

Teniendo en cuenta las expresiones (6)-(7) y dado que sabemos que $f\left(\frac{\tau}{y}\right) = \frac{f(y,\tau)}{f(y)}$, la recaudación por IRPF condicionada a una determinada renta (y) puede expresarse como

$$R(y) = \frac{1}{f(y)} \int_0^{\infty} \tau f(y, \tau) d\tau, \quad (9)$$

mientras que la recaudación total por IRPF se puede calcular a través de la siguiente expresión

$$R = \int_0^{\infty} \tau f(y, \tau) d\tau dy. \quad (10)$$

De este modo, en las ecuaciones (9)-(10) el lector puede apreciar cómo la recaudación por IRPF para un nivel de renta dado, así como la recaudación total por IRPF dependen de la recaudación por IRPF, de la distribución conjunta de las rentas salariales y la recaudación por IRPF, y de la distribución marginal de los salarios.

Teniendo en cuenta todo lo anterior, y dado que sabemos que $R = E[T(y)]$ y $g(y) = f(y)R(y)$, la recaudación fiscal esperada, \hat{R} , se puede calcular a través de la siguiente expresión

$$\hat{R} = \int_0^{\infty} \hat{T}_h(y)\hat{g}_h(y)dy, \quad (11)$$

donde $\hat{g}_h(y)$ es el estimador de la distribución de rendimientos fiscales para un nivel de renta dada y $\hat{T}_h(y)$ es el estimador de la estructura impositiva.

En este contexto, (Creedy, 2010), (Creedy y Gemmell, 2006) y (Onrubia y Picos, 2011), entre otros, proponen evaluar estos rendimientos fiscales a través de formas totalmente paramétricas tanto para la estructura impositiva, $T(y)$, como para la distribución de la renta, $g(y)$, que pueden no estar avaladas por los datos. Sin embargo, esta falta de robustez a errores de especificación puede provocar serios errores de cálculo en las elasticidades de los ingresos fiscales a cambios en la renta o en la estructura impositiva.

Para evitar este tipo de situaciones, en este proyecto proponemos calcular los rendimientos fiscales esperados a través de técnicas no paramétricas como las utilizadas en la sección anterior que no necesitan para su cálculo asumir previamente una forma funcional concreta para las funciones $T(y)$ y $g(y)$. De este modo, utilizando técnicas clásicas de estimación no paramétricas como el conocido estimador de Nadaraya-Watson (ver (Nadaraya, 1964) y (Watson, 1964)), proponemos calcular la relación entre los ingresos fiscales por IRPF (T) y las rentas salariales (Y) a través del siguiente estimador no paramétrico

$$\check{T}_h(y) = \frac{\frac{1}{Nh} \sum_{i=1}^N K\left(\frac{y-Y_i}{h}\right) \tau_i}{\check{f}_h(y)}, \quad (12)$$

donde $\check{f}_h(y)$ ya ha sido calculado previamente en la ecuación (1), mientras que el estimador de la distribución de rendimientos fiscales para un nivel de renta dada tiene la siguiente forma

$$\check{g}_h(y) = \check{f}_h(y) \check{R}_h(y) = \frac{1}{Nh} \sum_{i=1}^N K\left(\frac{y-Y_i}{h}\right) T_i. \quad (13)$$

Finalmente, reemplazando (12)-(13) en (11) y reordenando términos obtenemos el siguiente estimador no paramétrico para la recaudación fiscal esperada

$$\check{R} = \frac{1}{Nh} \sum_{i=1}^N \tau_i \int_0^{\infty} K\left(\frac{y-Y_i}{h}\right) dy. \quad (14)$$

En este punto cabe destacar que las expresiones (12)-(13) sólo pueden utilizarse de forma correcta cuando las variables aleatorias son continuas y siguen una distribución simétrica. Sin embargo, tanto la renta salarial como los ingresos fiscales sólo toman valores positivos y la información disponible está representada a través de intervalos. Con el objetivo de resolver esta situación, recurrimos de nuevo a la función kernel Gamma introducida previamente, que garantiza que los kernels empleados para el cálculo de \hat{R} nunca asignan pesos fuera del soporte, y al procedimiento desarrollado previamente para datos agrupados. Con todo ello, el estimador propuesto para calcular la estructura impositiva tiene la siguiente forma

$$\hat{T}_h(y) = \frac{\sum_{j=1}^J \rho_j \tau_j K_{j,h}(y)}{\hat{f}_h(y)}, \quad (15)$$

donde $\hat{f}_h(y)$ se calcula como en la ecuación (5), “j” representa los intervalos en los que se recoge la información de renta, \bar{l}_j es el punto medio del intervalo y ρ_j es el porcentaje de la población acumulada en cada intervalo. Por su parte, el estimador resultante para la recaudación fiscal esperada es el siguiente

$$\hat{R} = \sum_{j=1}^J \rho_j \tau_j \int_0^{\infty} K_{\bar{l}_j, h}(y) dy. \quad (16)$$

Finalmente, y tal y como se destaca en (De las Heras y Rodríguez-Poo, 2015), la elasticidad de los rendimientos fiscales a cambios en la distribución de la renta puede calcularse estimando el valor de la derivada de $T(y)$. En otras palabras, siendo $\alpha = T(y)$ y $\beta = T'(y)$, donde $T'(y)$ es la derivada de primer orden de $T(y)$, proponemos recurrir a un procedimiento de estimación no paramétrico local lineal (ver (Fan y Gijbels, 1996)) que equivale a minimizar la siguiente función criterio

$$\sum_{i=1}^N (\tau_i - \alpha - \beta(y - Y_i))^2 K\left(\frac{y - Y_i}{h}\right) \quad (17)$$

con respecto a α y β obteniendo

$$\begin{pmatrix} \check{\alpha}(y) \\ \check{\beta}(y) \end{pmatrix} = (\check{Y}^T W \check{Y})^{-1} \check{Y}^T W T, \quad (18)$$

donde W es una matriz diagonal $N \times N$ de pesos de la forma $W = \text{diag}\{K_h(y - Y_i)\}$ y \check{Y} es una matriz $N \times 2$ cuyo i th elemento es de la forma $\check{Y}_i = [1, (Y_i - y)]$.

Para evitar los problemas de soporte comentados previamente y tener en cuenta que la información se proporciona en intervalos, recurrimos de nuevo a la función kernel Gamma con lo que el estimador resultante es

$$\begin{pmatrix} \hat{\alpha}(y) \\ \hat{\beta}(y) \end{pmatrix} = (\check{Y}^T \bar{W} \check{Y})^{-1} \check{Y}^T \bar{W} T, \quad (19)$$

donde ahora $\bar{W} = \text{diag}\{\rho_j K_{\bar{l}_j, h}(y)\}$.

Revisando la literatura, no se encuentra estudio alguno que considere los estimadores propuestos en esta sección. Sin embargo, combinando las estructuras de prueba de (Chen, 2000) y (Reyes et al., 2019) es posible demostrar que los estimadores $\hat{T}_h(y)$, \hat{R} , $\hat{\alpha}(y)$ y $\hat{\beta}(y)$ exhiben buenas propiedades asintóticas de modo que estos estimadores convergen en probabilidad a sus verdaderos valores a medida que aumenta el tamaño muestral. Además, podemos realizar inferencias robustas dado que también se puede demostrar que, a medida que aumenta el tamaño muestral, la distribución de los estadísticos correspondientes puede aproximarse mediante una distribución con media cero y varianza a determinar. No obstante, el estudio concreto de estas propiedades queda fuera del objetivo de este proyecto y se deja pendiente para una investigación futura.

4.2.2. Elasticidad fiscal y desigualdad salarial

Con el objetivo de predecir el comportamiento de la recaudación fiscal a cambios en el ciclo económico, en esta sección nos centramos en establecer un marco teórico que permita determinar cómo los rendimientos fiscales, $R(y)$, presentan una cierta relación desconocida con algún parámetro que represente la desigualdad en la distribución de las rentas salariales como el índice de Gini.

Tomando como referencia la propuesta en (De las Heras y Rodríguez-Poo, 2015), suponemos que la distribución de la renta salarial, $f(y)$, pertenece a un conjunto de funciones parametrizables en términos del índice de Gini, θ , así como de otros parámetros que caracterizan dicha distribución, μ , i.e., $f(y; \theta, \mu)$. De este modo, estamos asumiendo que la función de distribución de la renta, $f(y)$, es conocida excepto por un determinado número de parámetros (θ, μ) .

En este contexto, es posible calcular la elasticidad de los ingresos fiscales por IRPF a cambios en la desigualdad salarial, para una renta salarial dada (y), a través de la siguiente expresión

$$\frac{\partial R(y)}{\partial \theta} = - \frac{\partial \log f(y; \theta, \mu)}{\partial \theta} \frac{g(y; \theta, \mu)}{f(y; \theta, \mu)}, \quad (20)$$

donde $g(y; \theta, \mu) = \int_0^\infty \tau f(y, \tau) d\tau = R(y; \theta, \mu) f(y; \theta, \mu)$. A este respecto cabe destacar que el signo de esta expresión indica cuál es el comportamiento esperado de los rendimientos fiscales ante un incremento en la desigualdad salarial. Si el signo es positivo podemos concluir que los rendimientos fiscales se incrementan a medida que aumenta la desigualdad, para un nivel de renta concreto (y), mientras que, si el signo es negativo, los rendimientos por IRPF disminuyen al incrementarse la desigualdad.

Con el objetivo de calcular esta elasticidad fiscal recurriremos a la información proporcionada por la AEAT en la “Estadística de los Declarantes por IRPF” que, a pesar de proporcionar la información en intervalos de renta y/o recaudación por IRPF, permite asociar la variable renta salarial (Y) con el Total de Ingresos computables o la variable de Ingresos fiscales por IRPF (T) con la Retención por rendimientos del trabajo. Desafortunadamente, la expresión (20) no puede ser calculada directamente con estos datos dado que depende de una serie de parámetros desconocidos a priori.

Para resolver esta situación, proponemos un procedimiento de estimación semiparamétrico de dos etapas alternativo que tenga en cuenta que las observaciones sobre la recaudación fiscal por IRPF (T) y las rentas salariales (Y) tienen solo soporte positivo y están agrupadas en “ j ” intervalos disjuntos ($A_j \times B_j$) que son compactos en \mathbb{R}^2 , donde en cada intervalo existe un total de n_j pares de renta y rendimientos fiscales, $\{(Y_{j1}, T_{j1}), \dots, (Y_{jn_j}, T_{jn_j})\}_{j=1}^J$, y donde $N = \sum_{j=1}^J n_j$. En concreto, en la primera etapa del procedimiento proponemos estimar los parámetros (θ, μ) utilizando técnicas clásicas de estimación máximo-verosímil, i.e.,

$$(\hat{\theta}, \hat{\mu}) = \min_{(\theta, \mu)} \sum_{j=1}^J \log f(Y_j; \theta, \mu), \quad (21)$$

mientras que en la segunda etapa nos centramos en la estimación no paramétrica de $g(y; \theta, \mu)$.

De este modo, utilizando los estimadores obtenidos minimizando (21) y recurriendo a un procedimiento no paramétrico estándar como el procedimiento local constante proponemos el siguiente estimador no paramétrico para $g(y; \theta, \mu)$,

$$\hat{g}_h(y; \hat{\theta}, \hat{\mu}) = \hat{R}_h(y; \hat{\theta}, \hat{\mu})f(y; \hat{\theta}, \hat{\mu}) = \sum_{i=1}^J \rho_j K_{\bar{l}_{j,h}}(y)T_j, \quad (22)$$

donde $\hat{R}_h(y; \theta, \mu)$ es el clásico estimador Nadaraya-Watson de la forma

$$\hat{R}_h(y; \hat{\theta}, \hat{\mu}) = \left(\frac{1}{f(y; \hat{\theta}, \hat{\mu})} \right) \sum_{j=1}^J \rho_j K_{\bar{l}_{j,h}}(y)T_j. \quad (23)$$

Finalmente, reemplazando (21)-(22) en (20) proponemos el siguiente estadístico para calcular la elasticidad de la recaudación fiscal por IRPF a cambios en el nivel de desigualdad

$$\frac{\partial \hat{R}_h(y)}{\partial \theta} = - \frac{\partial \log f(y; \hat{\theta}, \hat{\mu})}{\partial \theta} \frac{1}{f(y; \hat{\theta}, \hat{\mu})} \sum_{j=1}^J \rho_j K_{\bar{l}_{j,h}}(y)T_j \quad (24)$$

donde, asumiendo que $f(y; \theta, \mu)$ es la función de densidad de la distribución log-normal, la derivada de $\log f(y; \hat{\theta}, \hat{\mu})$ con respecto a la desigualdad, θ , se calcula del siguiente modo

$$\frac{\partial \log f(y; \hat{\theta}, \hat{\mu})}{\partial \theta} = \frac{\phi\left(\frac{1+\hat{\theta}}{2}\right)}{2\Phi\left(\frac{1+\hat{\theta}}{2}\right)} - \frac{(\log(y)-\hat{\mu})^2 \phi\left(\frac{1+\hat{\theta}}{2}\right)}{4\sqrt{2}}, \quad (25)$$

siendo $\phi(\cdot)$ y $\Phi(\cdot)$ la función de densidad y de distribución acumulada, respectivamente.

Destacar que el procedimiento de estimación propuesto en esta sección es de gran utilidad desde el punto de vista empírico dado que tiene en cuenta las principales características de estos datos (distribución claramente asimétrica con soporte solo positivo) evitando el sesgo asociado a los kernels simétricos y el hecho de que la información está suministrada a través de datos agrupados en intervalos. Al mismo tiempo, se introduce cierta flexibilidad en la especificación del modelo al suponer que la distribución de los rendimientos fiscales (i.e., $g(y)$) es una función totalmente desconocida que puede ser directamente estimada a través de los datos sin tener que asumir una forma paramétrica concreta para la misma. Finalmente, y siguiendo la estructura de prueba en (De las Heras y Rodríguez-Poo, 2015), es posible demostrar que el estadístico (24) presenta buenas propiedades asintóticas, es decir, a medida que el tamaño muestral aumenta $\frac{\partial \hat{R}_h(y)}{\partial \theta}$ es un estimador consistente de la elasticidad (20) y el estadístico $\sqrt{Nh} \left(\frac{\partial \hat{R}_h(y)}{\partial \theta} - \frac{\partial R(y)}{\partial \theta} \right)$ sigue una distribución normal con media cero y varianza finita. De este modo, es posible obtener intervalos de confianza de gran utilidad para realizar inferencias. Sin embargo, el estudio concreto de las propiedades asintóticas de este estimador queda fuera de los objetivos de este artículo y se deja pendiente para una investigación futura.

5. RESULTADOS EMPÍRICOS

En esta sección se muestran los resultados empíricos más importantes obtenidos utilizando la metodología propuesta en la Sección 4. En primer lugar, nos centraremos en el análisis de la distribución de las rentas salariales en España en los últimos años. Posteriormente, se analizará la elasticidad fiscal a cambios en la distribución de la renta y en la desigualdad salarial.

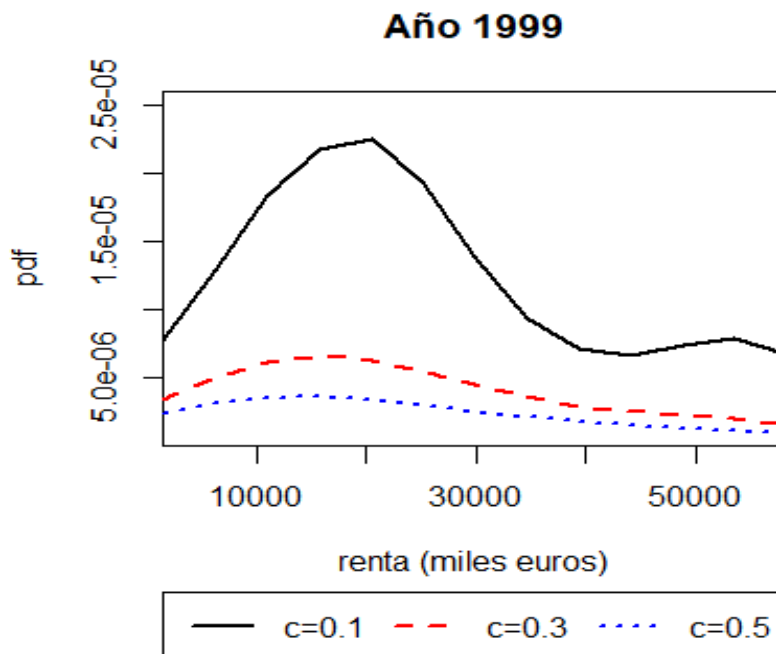
5.1. Distribución no paramétrica de las rentas salariales

Centrándonos en el estudio de la evolución de la distribución de las rentas salariales en España durante el periodo 1999-2020, recurrimos a los datos agrupados proporcionados por la estadística “Mercado de Trabajo y Pensiones en las Fuentes Tributarias” de la AEAT y detallados previamente en la Sección 3. En concreto, utilizaremos la información correspondiente a los tramos de salario, número de asalariados, distribución porcentual de los asalariados y salarios.

Con el objetivo de mostrar la sensibilidad de los estimadores de la función de densidad a la elección del ancho de ventana, en el Gráfico 1 recogemos la distribución de la renta estimada con el ancho de ventana $h = c\sigma J^{-2/13}$, donde $c=(0.1,0.3,0.5)$.

Gráfico 1

DISTRIBUCIÓN DE LOS SALARIOS EN EL AÑO 1999 CON DIFERENTES ANCHOS DE VENTANA

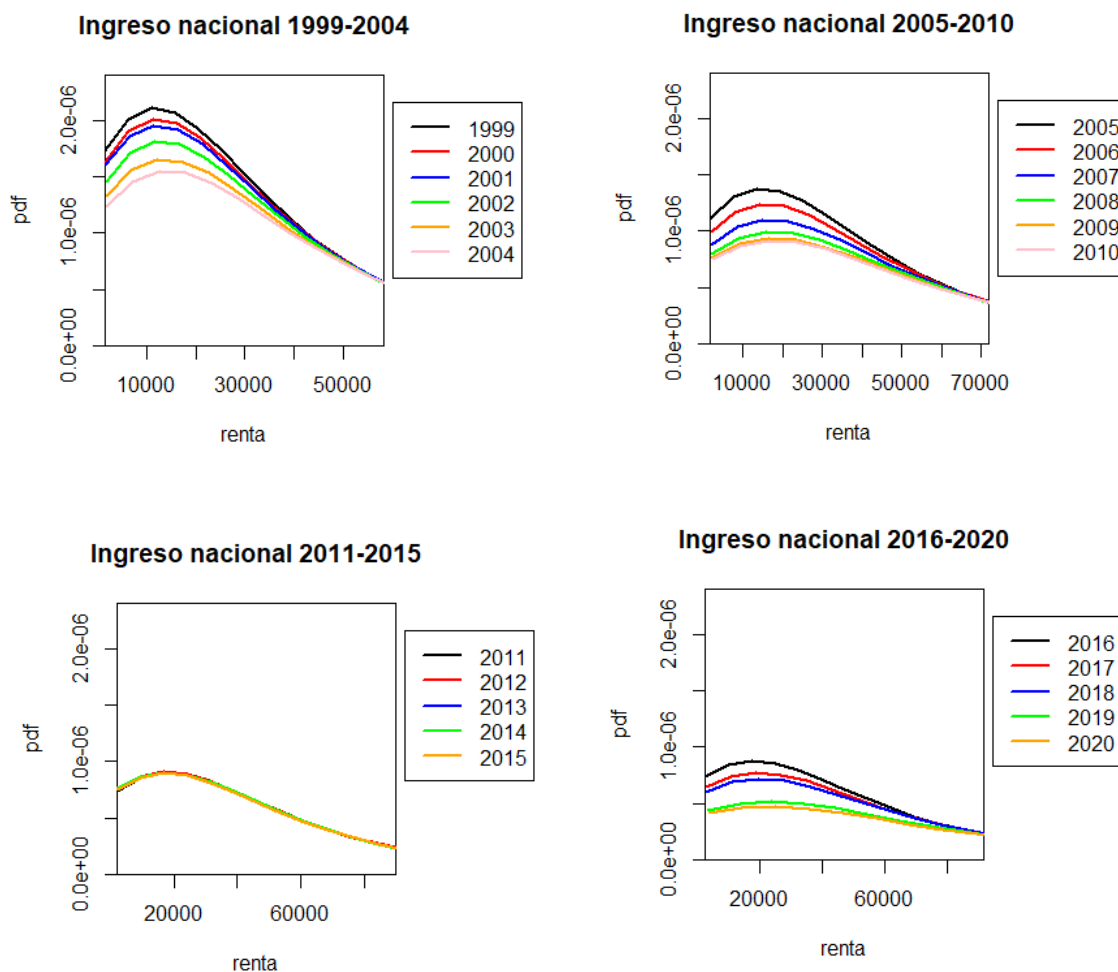


Fuente: elaboración propia con datos AEAT

Como era de esperar, el ancho de ventana juega un papel relevante en estos resultados dado que controla el grado de suavizado del estimador. Cuando $c=0.1$ las estimaciones están infrasuavizadas, mientras que cuanto mayor sea este parámetro más suavizada aparece la curva estimada.

En el Gráfico 2 aparecen representadas las estimaciones de la función de densidad de los salarios para el total nacional durante el periodo 1999-2020 utilizando el ancho de ventana óptimo $h = 0.78\sigma J^{-2/13}$.

Gráfico 2
DISTRIBUCIÓN DE LOS SALARIOS EN ESPAÑA (1999-2020)



Fuente: elaboración propia con datos AEAT.

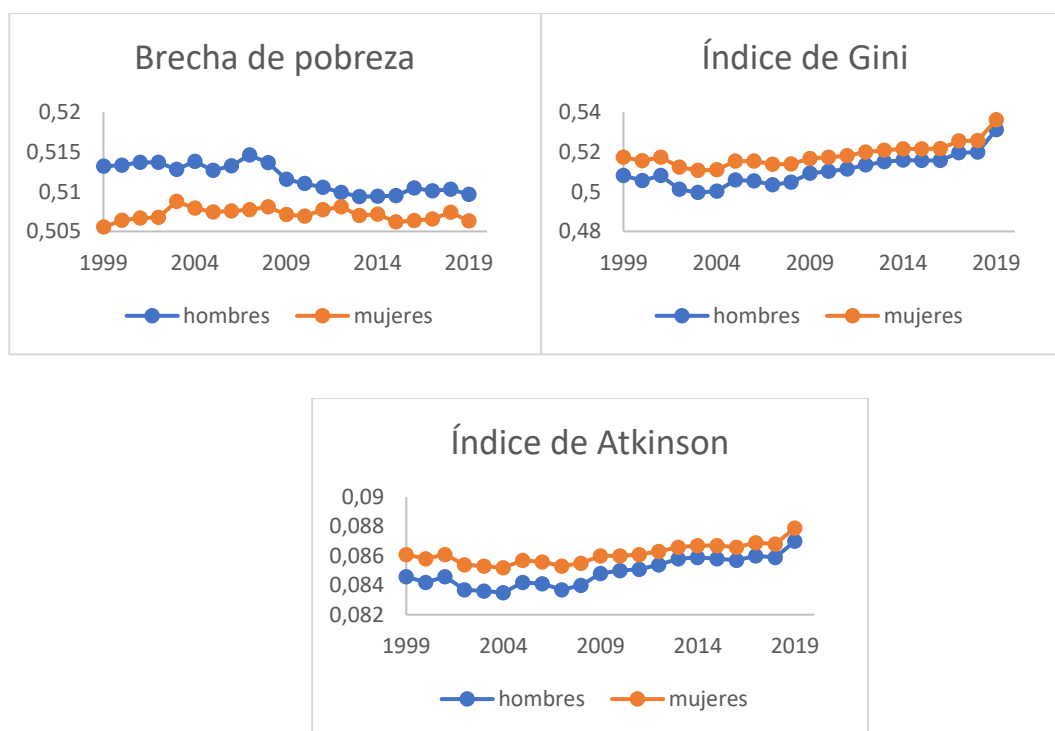
Las densidades salariales estimadas y recogidas en el Gráfico 2 se caracterizan por mostrar un comportamiento claramente asimétrico, especialmente en los primeros años del estudio, con una cola muy significativa a la derecha. Comparando su evolución a lo largo del tiempo se aprecia que los valores modales se han reducido considerablemente, lo que implica menores niveles de ingreso para la clase media y mayor desigualdad salarial. En concreto, estos resultados corroboran las cifras ofrecidas anualmente por Eurostat dado que se aprecia una disminución paulatina de los valores modales (o aumento de la desigualdad) hasta el año 2009, seguido por un periodo de estancamiento hasta el año 2016. De este modo, en términos relativos, podríamos concluir que la política de consolidación fiscal llevada a cabo por el gobierno español a partir del año 2010 parece haber conseguido contener el aumento de la desigualdad entre los individuos. Sin embargo, es

muy preocupante el aumento de la desigualdad que se aprecia en los años 2019 y 2020 y que se prevé que sea incluso mayor en los próximos años por el impacto de la crisis generada por la Cov-19.

Asimismo, estos resultados se ven corroborados por las distintas medidas de pobreza y desigualdad representadas en las figuras recogidas en el Gráfico 3. En concreto, para calcular la tasa de riesgo de pobreza seguimos las indicaciones establecidas por Eurostat y usamos el 60% de la mediana del salario de los hogares, mientras que fijamos el parámetro de aversión igual a 0.5 para estimar el índice de Atkinson. Además, dado el gran impacto que tienen los resultados del año 2020 en los distintos índices considerados y con el objetivo de poder analizar más claramente su evolución a lo largo de los años, en el Gráfico 3 se considera el periodo 1999-2019.

Gráfico 3

TASA DE RIESGO DE POBREZA, ÍNDICE DE GINI E ÍNDICE DE ATKINSON



Fuente: elaboración propia con datos AEAT.

El índice de Gini y de Atkinson recogidos en el Gráfico 3 confirman que se ha producido un aumento de la desigualdad generalizado durante todo el periodo de estudio tanto para hombres como para mujeres, aunque con distintas intensidades. Desde 1999 hasta 2004 se aprecia una reducción de la desigualdad, seguido por un periodo de bonanza económica que no afectó en gran medida a la desigualdad. Sin embargo, una vez desencadenada la crisis de 2007, se aprecia un incremento considerable hasta el año 2018, seguido por incremento exponencial a partir de entonces. En términos de la tasa de riesgo de pobreza se aprecia que ésta se ha mantenido más o menos estable hasta el año 2007 que experimenta un crecimiento exponencial pero que rápidamente es reducida hasta el final del periodo considerado.

Comparando la evolución entre hombres y mujeres en lo que respecta al índice de Gini y de Atkinson se aprecia una evolución similar para ambos grupos poblacionales, aunque la diferencia existente se ha ido reduciendo con el paso de los años. Finalmente, en lo que respecta a la tasa de riesgo de pobreza sí que se observa un proceso de convergencia mucho más claro que con los indicadores anteriores, aunque en los últimos años se ha visto que este indicador se ha visto incrementado en mayor medida en el grupo de los hombres. Un análisis pormenorizado de este efecto deberá ser desarrollado en los próximos años, una vez se disponga de información más detallada sobre los ingresos de los hogares durante ese periodo y sobre el efecto de la pandemia Cov-19.

5.2. Elasticidad fiscal a cambios en la distribución de la renta

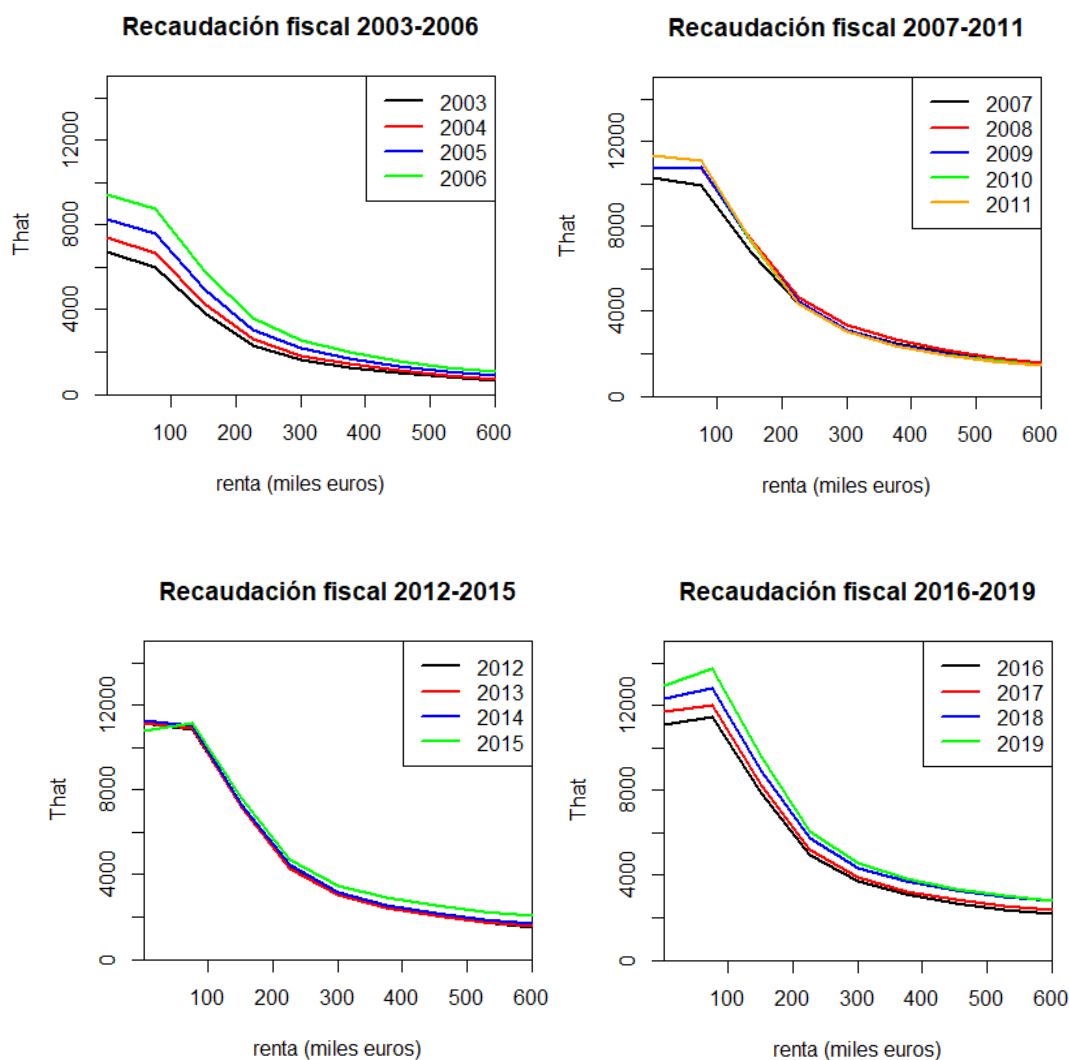
Con el objetivo de analizar la elasticidad de la distribución de las rentas salariales a cambios en la distribución de la renta o en el nivel de desigualdad recurrimos a los datos recogidos en la publicación “Estadística de los Declarantes del IRPF”. Como se ha comentado previamente, estos datos se caracterizan por estar recogidos en intervalos de renta asimétricos con lo que los procedimientos de estimación proporcionados en la Sección 4 son de gran utilidad. En primer lugar, nos centraremos en la realización de un análisis sobre la evolución de la recaudación fiscal en España. En segundo lugar, estudiaremos la evolución de la elasticidad de los ingresos fiscales por IRPF a cambios en el nivel de renta salarial.

Como se ha comentado en la Sección 3, estos datos están basados en las declaraciones del IRPF correspondientes al ejercicio fiscal de referencia de la estadística y recoge información detallada sobre las principales partidas contenidas en el modelo de declaración. En nuestro caso concreto, utilizaremos como proxy de la renta salarial el epígrafe de Total de ingresos computables que supone la suma de los siguientes conceptos: (i) Retribuciones dinerarias, (ii) Retribuciones en especie importe íntegro, (iii) Contribuciones empresariales a planes de pensiones, planes de previsión social empresarial y mutualidades de previsión social, (iv) Aportaciones al patrimonio protegido de las personas con discapacidad. Por su parte, con el objetivo de aproximar la información sobre los ingresos fiscales por IRPF utilizamos como proxy la variable denominada Retención por rendimientos del trabajo.

En el Gráfico 4 se recoge la evolución de la recaudación fiscal para el total nacional durante el periodo 2003-2019 utilizando el estimador propuesto en la ecuación (15) y utilizando el ancho de ventana óptimo señalado previamente (i.e., $h=0.78\sigma_J^{-2/13}$). En cada uno de los gráficos representamos en el eje vertical los valores estimados de $T(y)$ y en el eje horizontal la renta salarial. De este modo, los valores recogidos por estas gráficas nos permiten establecer la relación existente entre los ingresos fiscales por IRPF y los rendimientos salariales de los contribuyentes para el total nacional. Asimismo, seremos capaces de identificar aquellos hogares que están sometidos a una mayor carga impositiva y cómo ha sido su evolución a lo largo del tiempo.

Gráfico 4

ESTIMACIÓN DE LA RECAUDACIÓN FISCAL EN ESPAÑA (2003-2019)



Fuente: elaboración propia con datos AEAT.

Analizando la evolución de la recaudación fiscal en España a lo largo de los años se aprecia un comportamiento claramente diferenciado dependiendo del periodo temporal considerado. En primer lugar, centrándonos en el periodo 2003-2006 se aprecia un aumento considerable de los ingresos fiscales por IRPF para todos los niveles de renta. En segundo lugar, se observa un aumento considerable en la recaudación impositiva a partir del año 2007, especialmente relevante para los hogares cuyo nivel de renta es inferior a los 100.000€. Recordar que este periodo coincide con la crisis financiera durante la cual el gobierno español aumentó diversos impuestos, entre ellos el IRPF, con el objetivo de poder hacer frente a la caída en la recaudación fiscal y a los elevados gastos asociados al estado de bienestar. A partir del año 2010 la recaudación impositiva prácticamente no ha sufrido variaciones hasta el año 2015 donde se aprecia que hay un ligero aumento de la recaudación fiscal para los hogares con renta superior. Finalmente, a lo largo del periodo 2016-2019 hay un aumento progresivo en la recaudación por IRPF generalizado para todos los

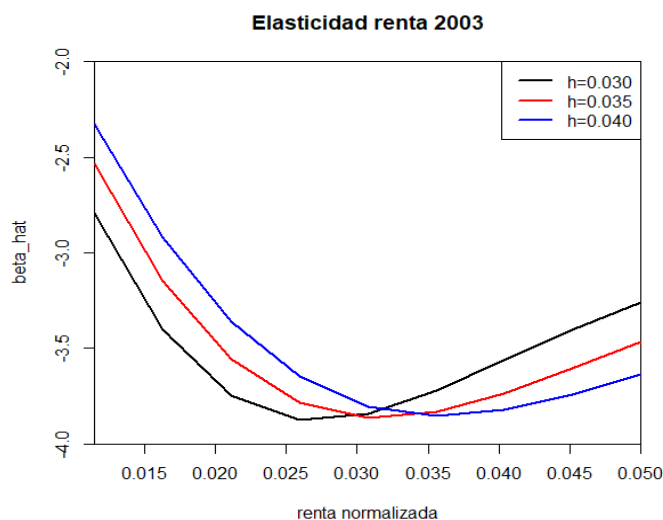
niveles de renta, aunque con un impacto considerablemente mayor para los hogares con menores ingresos y especialmente relevante en el año 2019.

Centrándonos en la elasticidad de la recaudación fiscal por IRPF a cambios en la distribución de la renta utilizamos los mismos datos que en el caso anterior y recurrimos a la técnica de estimación propuesta en la ecuación (19) centrándonos en los resultados obtenidos para el estimador $\hat{\beta}(y)$. Cabe destacar que a efectos de normalización hemos estandarizado en todos los casos las variables *Total de Ingresos computables* y *Retenciones del trabajo*.

Con el objetivo de analizar la sensibilidad de los estimadores propuestos al ancho de ventana, en el Gráfico 5 se recoge la estimación de la elasticidad fiscal a cambios en la distribución de la renta usando como ancho de ventana $h=(0.030,0.035,0.040)$. Posteriormente, en el Gráfico 6 se recogen estos estimadores para el periodo 2003-2019 utilizando como ancho de ventana $h=0.030$. En ambos casos en el eje vertical se representan los valores estimados de la primera derivada de la función asociada a la recaudación fiscal para un nivel de renta dado (eje horizontal).

Gráfico 5

ELASTICIDAD FISCAL A CAMBIOS EN LA DISTRIBUCIÓN DE LA RENTA EN EL AÑO 2003 CON DIFERENTES ANCHOS DE VENTANA

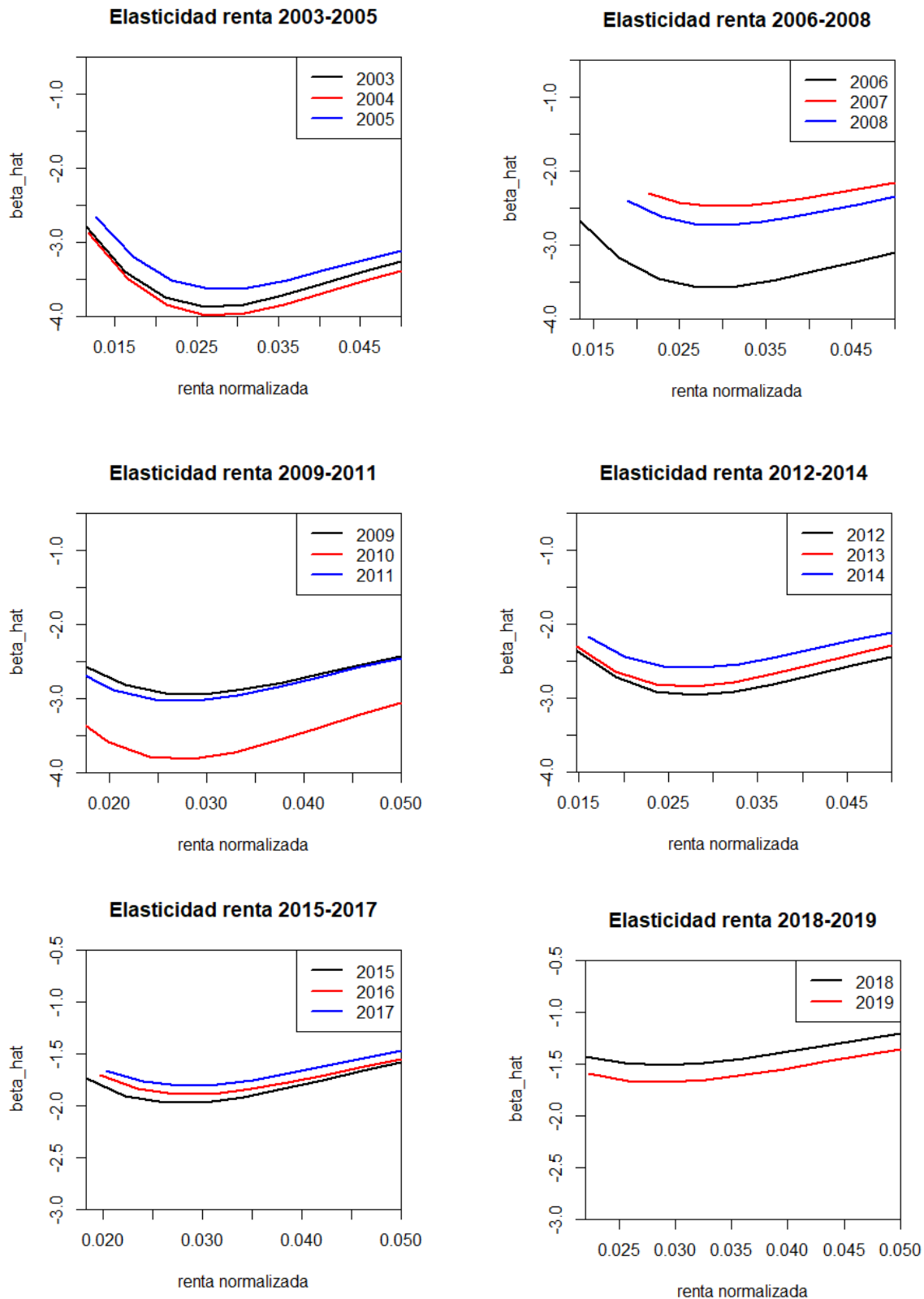


Fuente: elaboración propia con datos AEAT.

Como se aprecia en el Gráfico 5, el valor del ancho de ventana tiene un papel relevante a la hora de estimar la elasticidad fiscal a cambios en la distribución de la renta. Sin embargo, a diferencia de lo que ocurre en el Gráfico 1, valores más elevados del ancho de ventana no están asociados a un mayor grado de suavizado. En este caso concreto, se aprecia un desplazamiento hacia la derecha del valor mínimo de la curva estimada a medida que aumenta (h), es decir, del valor a partir del cual incrementos en los niveles de renta están asociados a aumentos en la recaudación por IRPF. Este resultado refleja el papel relevante que juega el ancho de ventana en este tipo de estimadores y un análisis más pormenorizado de las propiedades asintóticas de este estimador sería conveniente con el objetivo de obtener un ancho de ventana óptimo como el que se propone en la Sección 4.1 para el estimador de la función de densidad.

Gráfico 6

ELASTICIDAD FISCAL A CAMBIOS EN LA DISTRIBUCIÓN DE LA RENTA (2003-2019)



Fuente: elaboración propia con datos AEAT.

A la vista de los resultados obtenidos en el Gráfico 6 las estimaciones obtenidas reflejan una relación significativa entre la recaudación fiscal y la renta salarial, tal y como cabría esperar dada la estructura de este impuesto, dado que las curvas estimadas varían de forma estable alrededor del mismo valor mínimo. En términos generales, se aprecia una elasticidad recaudatoria claramente negativa en los tramos de renta que representan más del 80% de las retribuciones salariales.

Además, si analizamos el comportamiento de esta elasticidad a lo largo de los años se aprecia que éste no es homogéneo. En el periodo 2003-2006 prácticamente no se aprecian diferencias. Sin embargo, en el periodo 2007-2010, coincidiendo con un periodo de fuerte crisis económica y financiera, se puede observar una mayor elasticidad fiscal, especialmente relevante para los hogares con mayor nivel de renta. En concreto, se produce un desplazamiento hacia la derecha del punto de inflexión de la curva estimada y los hogares por debajo de una renta normalizada con valor 0.017 desaparecen. Posteriormente, a partir del año 2011 el punto de inflexión de la curva vuelve a reducirse y los hogares con menores rentas salariales vuelven a tributar por IRPF.

Finalmente, destacar que la forma convexa que exhibe la curva estimada de esta elasticidad al comienzo del estudio se ve amortiguada considerablemente con el paso de los años, siendo casi plana a partir del año 2017. Este resultado es un claro indicador de que el esfuerzo fiscal que están haciendo los hogares en los últimos años para mantener los recursos públicos es bastante homogéneo para todos los niveles de renta.

5.3. Elasticidad fiscal a cambios en la desigualdad salarial

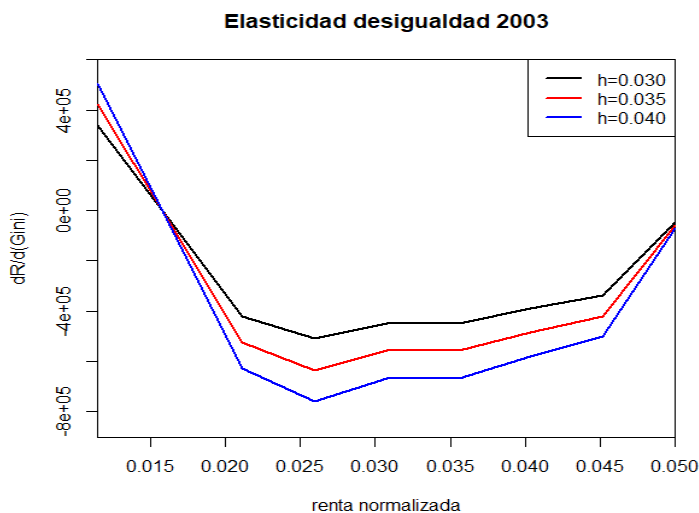
En esta sección nos centraremos en analizar la relación existente entre la recaudación fiscal y la desigualdad salarial de los contribuyentes españoles para el periodo comprendido entre 2003 y 2019. Para ello, utilizaremos la base de datos obtenida de la “Estadística de Declarantes del IRPF” de la AEAT comentada previamente y las variables utilizadas son el *Total de Ingresos computables* como proxy para la renta salarial y la *Retención por rendimientos del trabajo* como proxy de los ingresos fiscales por IRPF. En ambos casos estas variables son normalizadas para facilitar la interpretación de los resultados.

Teniendo en cuenta esta información, hemos procedido a calcular los valores de $\frac{\partial R(y)}{\partial \theta}$ para cada año disponible y para diferentes valores del rendimiento del trabajo utilizando la técnica de estimación descrita en la Sección 4.2.2. Asimismo, con el objetivo de mostrar la sensibilidad de esta estimación a la elección del ancho de ventana, en el Gráfico 7 se muestran las curvas estimadas usando como ancho de ventana los valores $h=(0.030,0.035,0.040)$. Posteriormente, en el Gráfico se recogen las curvas estimadas cuando $h=0.03$ para el periodo de estudio considerado, 2003-2019. En ambos casos el eje vertical representa los valores estimados de $\frac{\partial R(y)}{\partial \theta}$, mientras que el eje horizontal recoge los rendimientos del trabajo.

De este modo, los Gráficos 7 y 8 miden la variación de los ingresos fiscales por IRPF cuando varía el nivel de desigualdad (θ) para un nivel de renta dado (y), desde el punto de vista económico,

estos resultados se pueden interpretar como la elasticidad de la recaudación fiscal por IRPF a cambios en la desigualdad salarial.

Gráfico 7
ELASTICIDAD FISCAL A CAMBIOS EN LA DESIGUALDAD CON DIFERENTE ANCHO DE VENTANA (AÑO 2003)

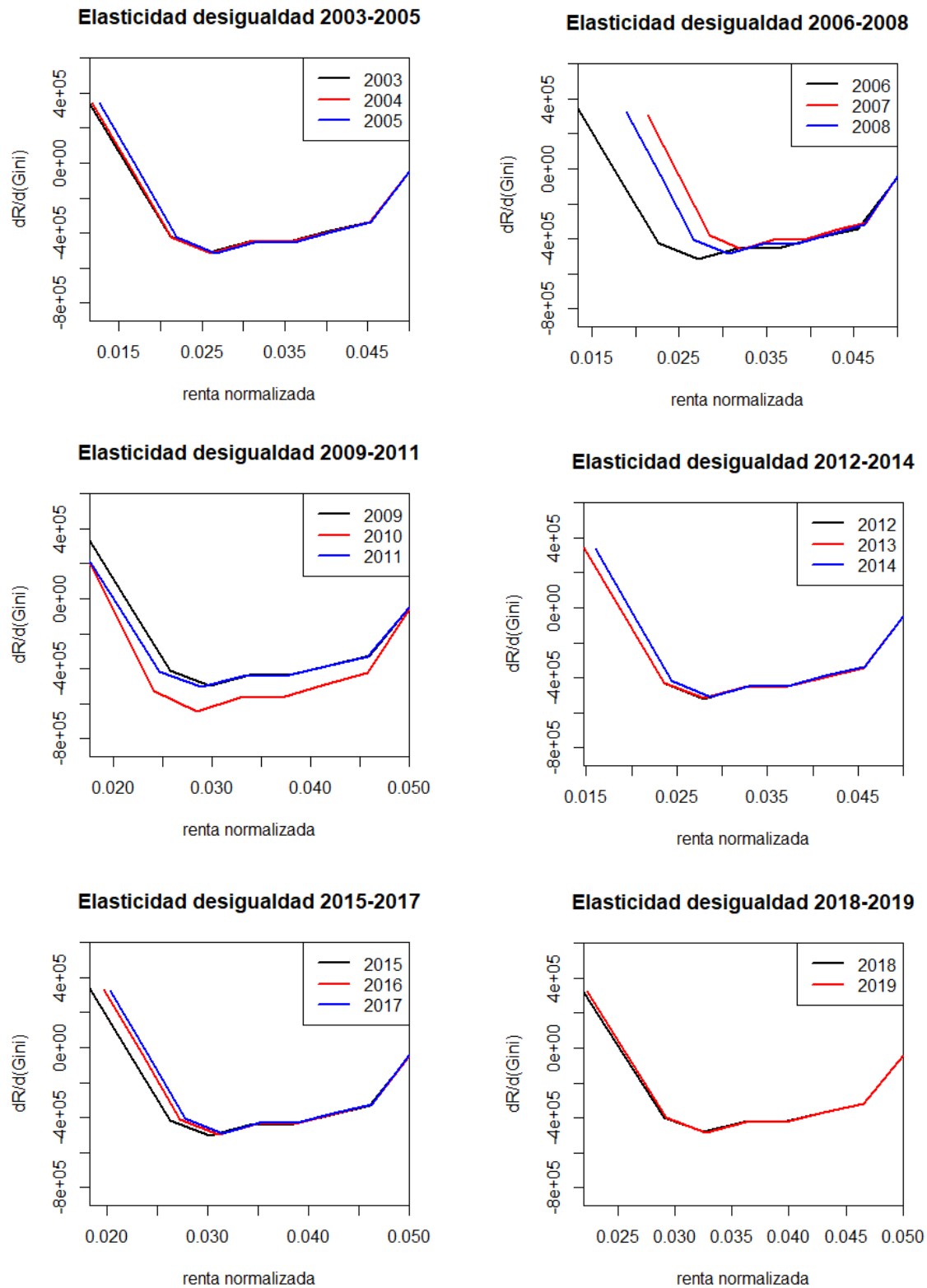


Fuente: elaboración propia con datos AEAT.

Analizando los resultados del Gráfico 7 se aprecia claramente el papel relevante que ocupa la elección del ancho de ventana en este tipo de estimadores no paramétricos. A mayor nivel de (h) menor es el valor del estimador de la elasticidad fiscal a cambios en la desigualdad salarial, aunque en todos los casos analizados las curvas estimadas oscilan alrededor del mismo punto de inflexión de la renta salarial. De este modo, al igual que ocurre en el Gráfico 5 y a la vista de estos resultados, sería conveniente realizar un estudio específico sobre cuál es el ancho de ventana óptimo para este estimador concreto.

Gráfico 8

ELASTICIDAD FISCAL A CAMBIOS EN LA DESIGUALDAD (2003-2019)



Fuente: elaboración propia con datos AEAT.

Analizando los resultados del Gráfico 8 se aprecia que la elasticidad recaudatoria es negativa en los tramos centrales de la distribución de rentas salariales, aunque con cierto nivel de desplazamiento del punto de inflexión de las curvas estimadas a lo largo del periodo de estudio. Este es un resultado relevante dado que nos permite identificar cómo los distintos ciclos económicos que ha experimentado la economía española a lo largo de los años han impactado sobre la elasticidad de la recaudación fiscal ante modificaciones en la desigualdad salarial.

En concreto, a lo largo del periodo de bonanza económica, 2003-2006, prácticamente no se producen oscilaciones en las curvas estimadas y se observa una relación negativa entre la recaudación fiscal y la desigualdad salarial en los niveles próximos a la media de los ingresos salariales. Sin embargo, en el año 2007, coincidiendo con el inicio de la crisis financiera, se aprecia un claro desplazamiento hacia la derecha del punto de inflexión a partir del cual esta elasticidad estimada pasa a ser positiva. Este resultado es un claro reflejo de la caída recaudatoria que se produjo durante los años 2007-2009 al aumentarse la desigualdad salarial. Posteriormente, a partir del año 2009 y de manera continuada hasta el año 2015 se aprecia un desplazamiento del punto de inflexión de las curvas estimadas hacia la izquierda de modo que podemos concluir que durante este periodo hay un incremento en la recaudación fiscal por IRPF dado que la obtención de valores positivos se produce en rentas salariales más bajas que en los periodos anteriores. Finalmente, a partir del año 2016 se aprecia que el comportamiento de la elasticidad fiscal ante cambios en la desigualdad salarial permanece más o menos homogéneo a lo largo de los años. Con todo ello, queda demostrado que, a pesar de que España es uno de los países que menos ha reducido la desigualdad salarial en los últimos años, ha conseguido que la capacidad recaudatoria del IRPF no recaiga en los hogares más vulnerables.

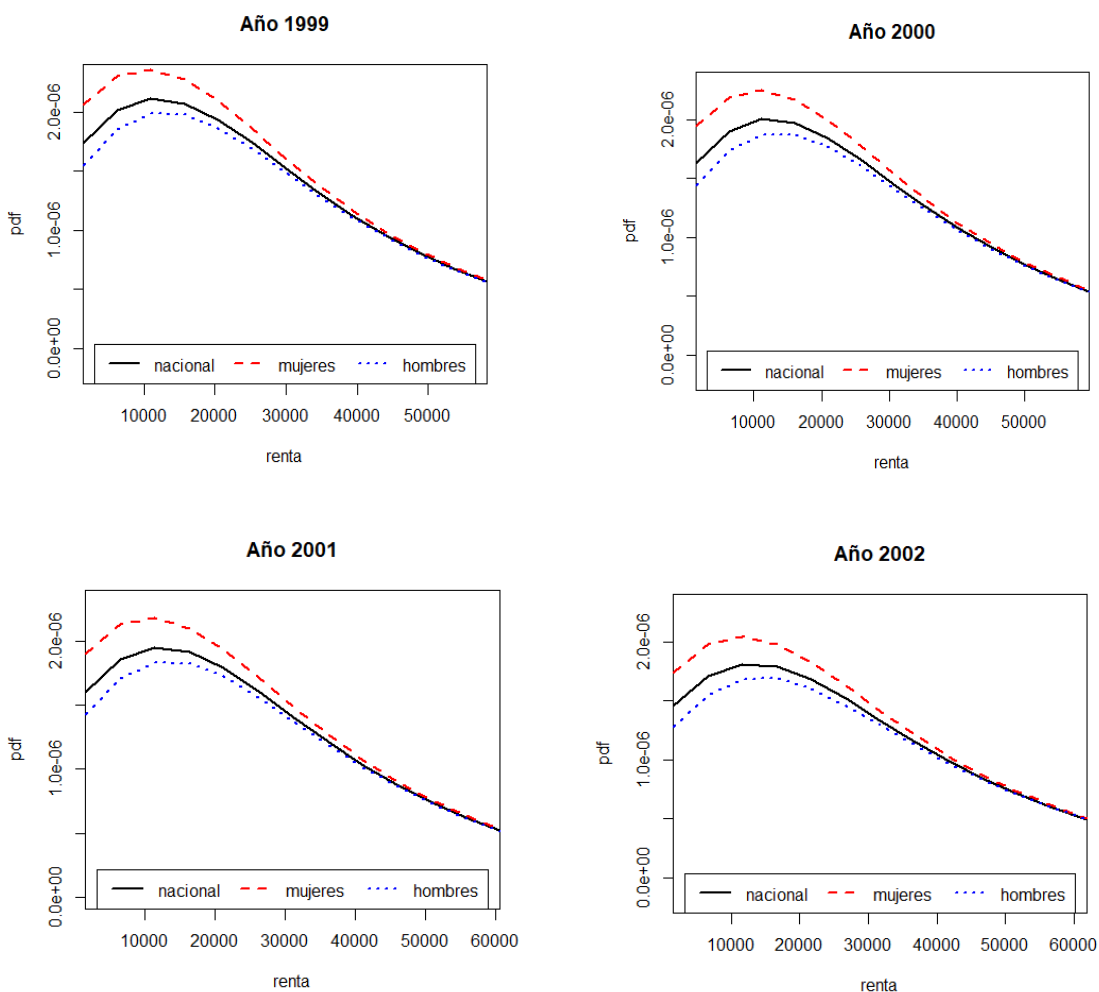
6. CONCLUSIONES

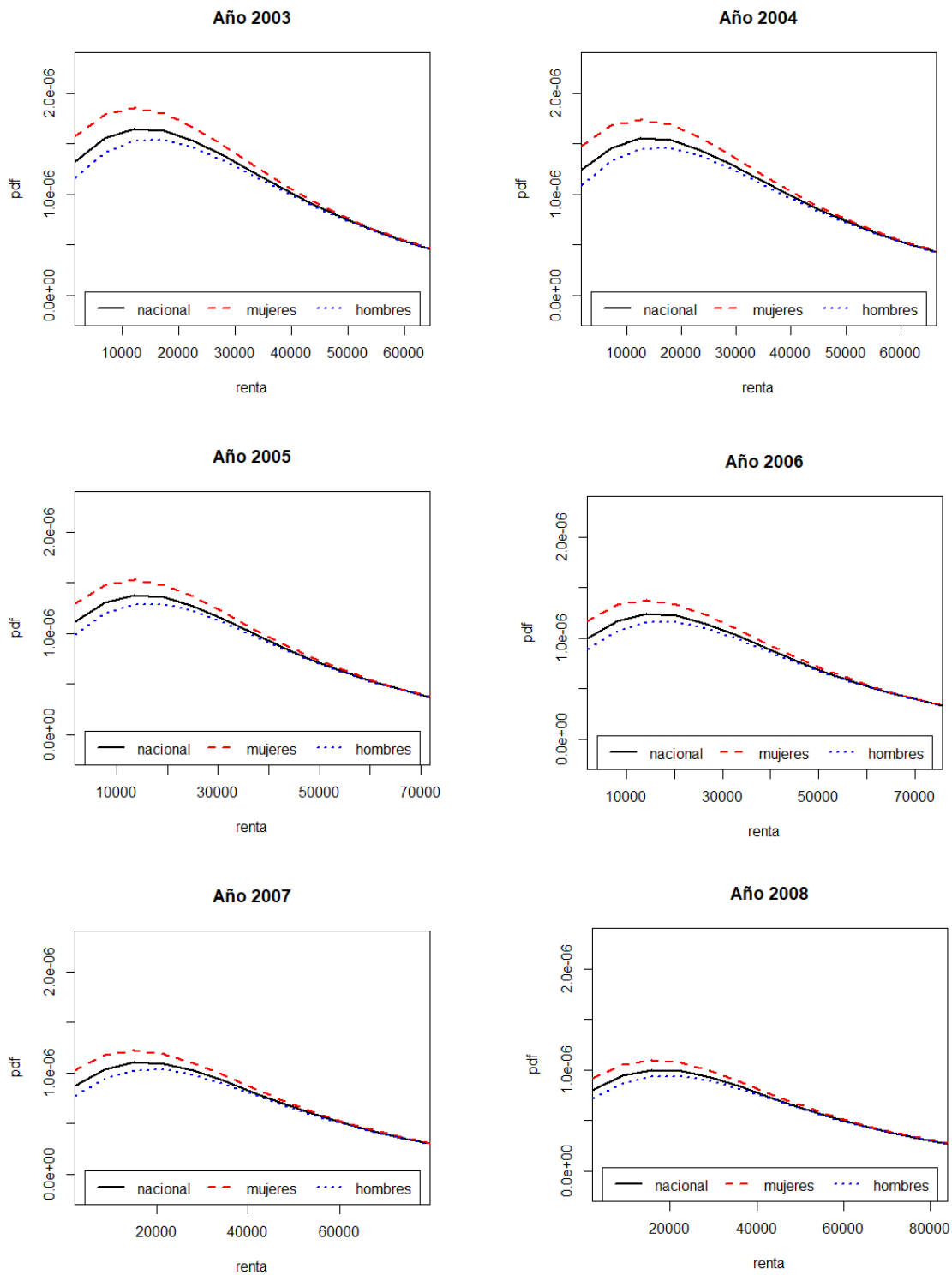
En este artículo proponemos la utilización de la información recogida en distintos registros administrativos de la AEAT para analizar el comportamiento de los ingresos fiscales por IRPF ante variaciones en las rentas salariales y en la desigualdad salarial en España. Dado que la principal característica de estos datos es que están agrupados en intervalos asimétricos y que las variables de interés solamente tienen soporte positivo, la mayoría de procedimientos de estimación tradicionales conducen a la obtención de estimadores inconsistentes. Con el objetivo de resolver esta situación, los objetivos de este trabajo se estructuran en dos grandes bloques. En un primer bloque, se propone una técnica de estimación no paramétrica que tiene en cuenta estas particularidades de los datos con el objetivo de estimar la función de densidad de los salarios (curva de Lorenz). En un segundo bloque, se establece un marco teórico no paramétrico con el objetivo de analizar la evolución esperada de la recaudación fiscal por IRPF ante variaciones en las rentas salariales y en los niveles de desigualdad de los contribuyentes españoles. Dado que los datos utilizados adolecen de la misma problemática que en objetivo anterior, se desarrollan distintos procedimientos de estimación no paramétrica que permiten corregir esta situación y obtener estimadores consistentes tanto de la elasticidad fiscal a variaciones en el nivel salarial como de la elasticidad de la recaudación por IRPF a cambios en la desigualdad salarial. Finalmente, analizando los resultados empíricos obtenidos con estos nuevos estimadores queda constatado que el

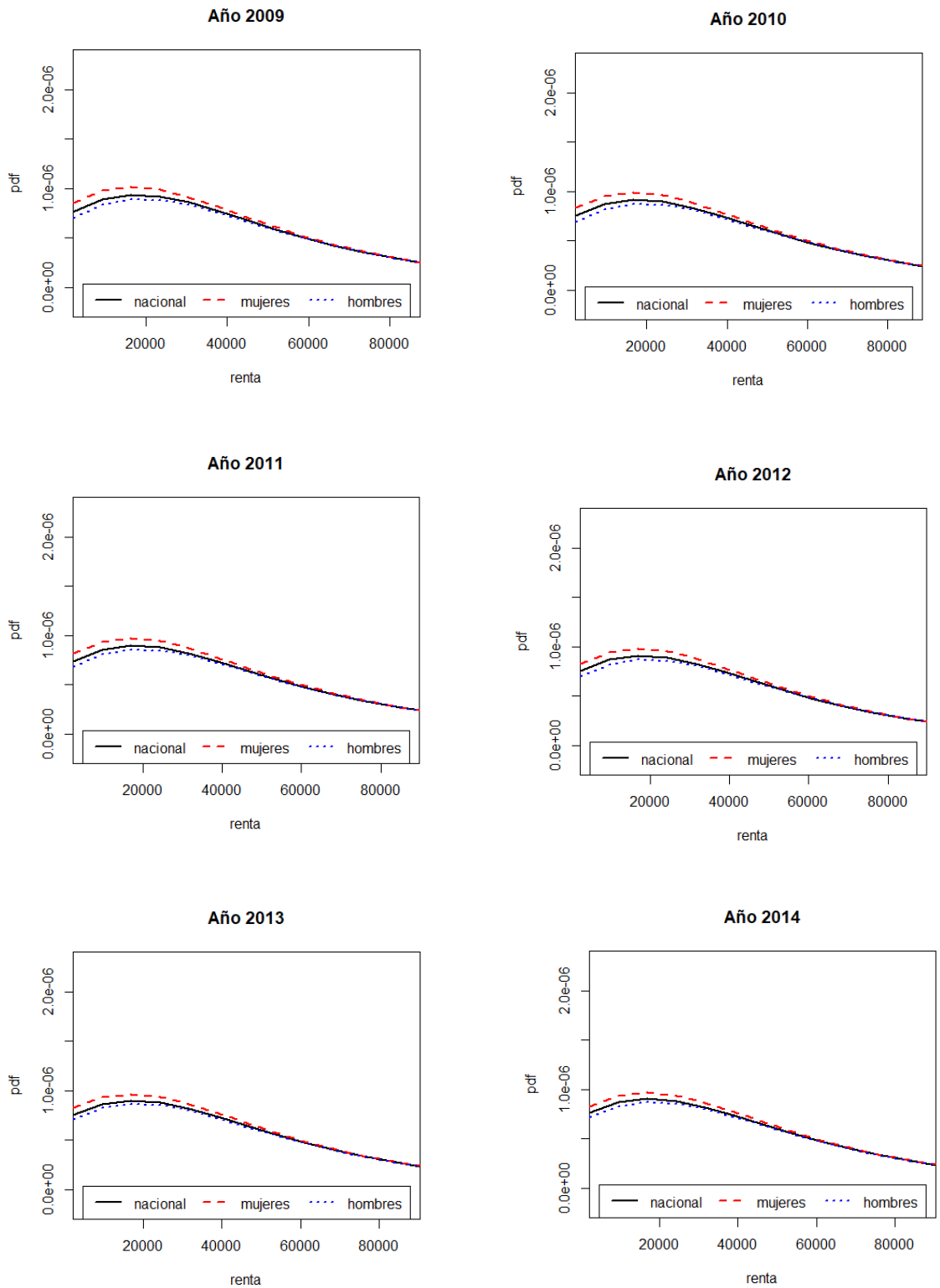
incremento de la desigualdad salarial no se ha producido en todos los grupos poblacionales de manera homogénea a lo largo de los años, sino que han sido los hogares con rentas más bajas los que han sufrido un mayor descenso en sus ingresos, especialmente en la última década. Asimismo, la elasticidad recaudatoria a cambios en la desigualdad es negativa a lo largo del periodo 2003-2019, mientras que la elasticidad fiscal a cambios en las retribuciones salariales exhibe un comportamiento claramente heterogéneo que permite identificar el impacto de las distintas crisis económicas sobre la capacidad recaudatoria por IRPF del Estado.

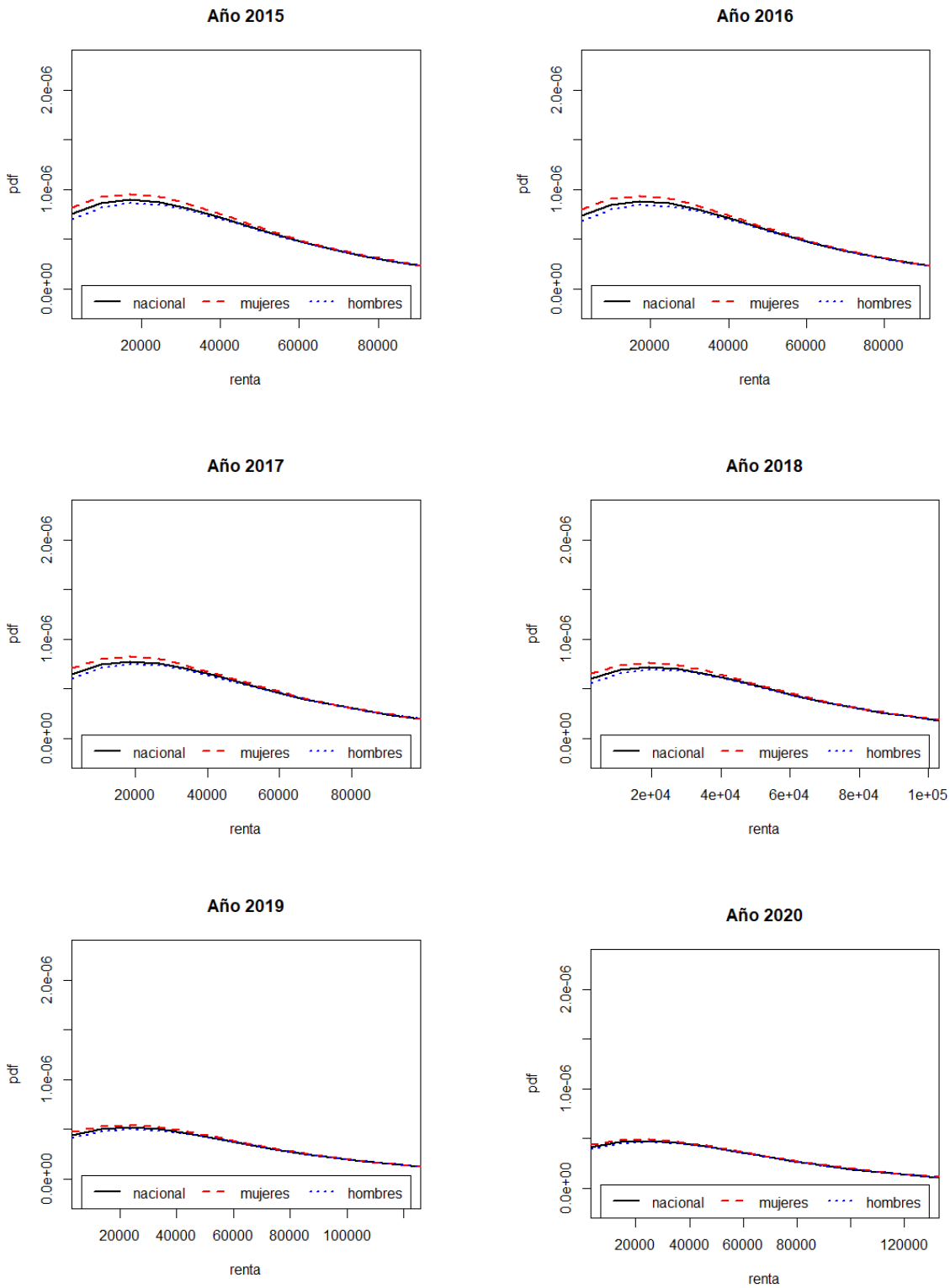
ANEXO

Con el objetivo de poder realizar un análisis comparativo en profundidad respecto al comportamiento de la distribución de los salarios entre hombres y mujeres y su evolución a lo largo del tiempo, en este Anexo se recogen los gráficos correspondientes a la distribución de los salarios desde el año 1999 hasta 2019.









Bibliografía

- ALESINA, A. y RODRIK, D. (1994): "Distributive politics and economic growth", *Quarterly Journal of Economics*, 109, pp. 465-490.
- ANAND, S. y SEGAL, P. (2008): "What do we know about global income inequality?", *Journal of Economic Literature*, Volumen 46, pp. 57-94.
- ANTONCZYK, D., FITZENBERGER, B. y SOMMERFELD, K. (2010): "Rising wage inequality, the decline of collective bargaining and the gender wage gap", *Labour Economics*, 17 (5), pp. 794-826.
- ARELLANO, M., BOVER, O. y BENTOLILA, S. (2002): "Unemployment duration, benefit duration and the business cycle", *Economic Journal*, 112 (479), pp. 223-265.
- ARRANZ, J., DAVIA, M., y GARCÍA-SERRANO, C. (2015): "Desigualdad salarial y empleo de bajos salarios en los países desarrollados. *Ekonomiaz*, nº 87, 1º semestre, pp. 60-87.
- ATKINSON, A. (2008) *The changing distribution of earnings in OECD countries*. Nueva York: Oxford University Press.
- AYALA, L. (2013): "Crisis económica y desigualdad", *Anuario de Relaciones Laborales en España*, 4, pp. 36-38.
- AYALA, L., MARTINEZ, R. y RUIZ HUERTA, J. (2013): *Desigualdad y redistribución en los países de la OCDE*. Madrid: Fundación Alternativas.
- BANDOURIAN, R., McDONALD, J. y TURLEY, R. (2003): "A comparison of parametric models of income distribution across countries and over time", *Estadística*, Volumen 55, pp. 135-152.
- BLAU, F. y KAHN, L. (1999): "Institutions and laws in the labor market", en *Handbook of Labor Economics*. Amsterdam: North-Holland, pp. 1399-1461.
- BLOWER, G. y KELSALL, J. (2002) "Nonlinear kernel density estimation for binned data: convergence in entropy", *Bernoulli*, Volumen 8, pp. 423-449.
- BLUNDELL, R., DUNCAN, A. y MEGHIR, C. (1998): "Estimating labor supply responses using tax reforms", *Econometrica*, 66(4), pp. 827-861.
- BLUNDELL, R., MEGHIR, C., SYMONS, E. y WALKER, I. (1988): "Labour supply specifications and the evaluation of tax reforms", *Journal of Public Economics*, Volumen 36, pp. 23-52.
- BONHOMME, S. y HOSPIDO, L. (2012a): *Earnings inequality in Spain: new evidence using tax data*, Banco de España: Mimeo.
- BONHOMME, S. y HOSPIDO, L. (2012b) *Earnings inequality in Spain: evidence from Social Security data*, Banco de España: Mimeo.
- BONHOMME, S. y HOSPIDO, L. (2017): The Cycle of Earnings Inequality: Evidence from Spanish Social Security Data. *Economic Journal*, 127 (603), pp. 1244-1278.
- BOURGUIGNON, F. y MAGNAC, T. (1990): "Labor supply and taxation in France", *Journal of Human Resources*, 25(3), pp. 358-389.
- BOURGUIGNON, F. y SPADARO, A. (2005): Microsimulation as a tool for evaluating redistribution policies. *Journal of Economic Inequality*, 4(1), pp. 77-106.
- BOUSTAN, L., FERREIRA, F., WINKLER, H. y ZOLT, E. (2013): "The effect of rising income inequality on taxation and public expenditures: evidence from US municipalities and school districts, 1970-2000", *The Review of Economics and Statistics*, 95(4), pp. 1291-1302.

- BROWN, L. y otros (2010): "The root-unroot algorithm for density estimation as implemented via waved block thresholding", *Probability Theory and Related Fields*, Volumen 146, pp. 401-433.
- CARRASCO, R., JIMENO, J. y ORTEGA, C. (2011): "Accounting for changes in the Spanish wage distribution: the role of employment composition", *Documento de Trabajo*, n.º. 1120, Banco de España.
- CARRASCO, R., JIMENO, J. y ORTEGA, C. (2015) "Returns to skills and the distribution of wages: Spain 1995-2010", *Oxford Bulletin of Economics and Statistics*, 77(4), pp. 542-565.
- CHEN, S. (2000): "Probability density function estimation using gamma kernels", *Annals of the Institute of Mathematical Statistical*, Volumen 52, pp. 471-480.
- CHETTY, R. (2009): "Is the taxable income elasticity sufficient to calculate deadweight loss? The implications of evasions and avoidance", *American Economic Journal: Economic Policy*, Volumen 1, pp. 31-52.
- CREEDY, J. (2009): *The elasticity of taxable income: an introduction*, s.l.: Research Paper Number 1085. Department of Economics. The University of Melbourne.
- CREEDY, J. (2010): *Income tax revenue: some simple analytics*, s.l.: Research Paper Number 1101. Department of Economics. The University of Melbourne.
- CREEDY, J. y GEMMELL, N. (2006): *Modelling Tax Revenue Growth*. s.l.:Cheltenham: Edward Elgar.
- CREEDY, J. y GEMMEL, N. (2013): "Measuring tax revenue responses to tax rates changes in multirate income tax systems: behavioral and structural factors". *International Tax and Public Finance*, 20(6), pp. 974-991.
- DE LAS HERAS, A., MURILLO, C. y RODRIGUEZ-POO, J. (2004): "Estimación de los efectos redistributivos y de las ganancias en bienestar social derivados de la progresividad del IRPF en las Comunidades Autónomas del Territorio del Régimen Común", *Cuadernos Económicos del ICE*, pp. 175-194.
- DE LAS HERAS, A. y RODRIGUEZ-POO, J. (2015): *Recaudación fiscal por IRPF y desigualdad salarial: un análisis empírico*, s.l.: Instituto de Estudios Fiscales.
- DÍAZ DE SARRALDE, S. y otros (2006): "La reforma del IRPF de 2007: una evaluación de sus efectos", *Papeles de Trabajo del Instituto de Estudios Fiscales*, PT 26/06.
- EUROSTAT (2015): *Income and Living Conditions*. [En línea] Disponible en: <http://ec.europa.eu/eurostat/web/income-and-living-conditions>
- FAN, J. y GIJBELS, I. (1996): "Local Polynomial Modelling and Its Applications", en: *Monographs on Statistics and Applied Probability* 66. s.l.:Chapman y Hall/CRC Monographs on Statistics y Applied Probability.
- FELDSTEIN, M. (1995): "The effect of marginal tax rates on taxable income: a panelstudy of the 1986 tax reform act", *Journal of Political Economy*, 103(3), pp. 551-572.
- FELDSTEIN, M. (1999): "Tax avoidance and the deadweight loss of the income tax", *The Review of Economics and Statistics*, 81(4), pp. 674-680.
- FUENTES, R. (2005): *Poverty, pro-poor growth and simulated inequality reduction*, s.l.: Human development report office occasional paper n.º 11.
- GAN, L. y TAN, J. (2011): *How does the optimal two-bracket income tax depend on wage inequality*, s.l.: Shanghai University of Finance and Economics. Discussion Paper.
- GARCÍA-PEREA, P. (1991): "Evolución de la situación salarial española desde 1963", en: S. Bentolila y L. Toharia, eds. *Estudios de Economía del Trabajo en España III: El problema del paro*. España: Ministerio de Trabajo y Seguridad Social, pp. 1289-1326.

- GARCÍA-SERRANO, C. y ARRANZ, J. (2013): "Crisis económica y desigualdad salarial", *Papeles de Economía Española*, n° 135, pp. 68-82.
- GIERTZ, S. (2009): "The elasticity of taxable income: influences on economy efficiency and tax revenue, and implications for tax policy", en: A. D. Viard, ed. *Tax Policy Lessons for the 2000s*. Washington D.C.: A.E.I. Press.
- GROSSMAN, H. (1991): "A general equilibrium model of insurrections", *American Economic Review*, 81, pp. 912-921.
- GRUBER, J. y SÁEZ, E. (2002): "The elasticity of taxable income: evidence and implications", *Journal of Public Economics*, 84(1), pp. 1-32.
- HAUSMAN, J. (1980): "The effect of wages, taxes and fixed costs on women's labor force participation", *Journal of Public Economics*, 14(2), pp. 161-194.
- HAUSMAN, J. y RUUD, P. (1984): "Family labor supply with taxes", *The American Economic Review*, 74(2), pp. 242-248.
- HENDERSON, D. y PARMETER, C. (2015): *Applied Nonparametric Econometrics*. New York: Cambridge University Press.
- HIDALGO, M. (2008): *Wage inequality in Spain 1980-2000*, s.l.: Universidad Pablo de Olavide. Economics Department, WP 08.08.
- HIDALGO, M. (2010): "Wage inequality in Spain, 1980-2000: the case of male head-of-household", *Estadística Española*, 52(174), pp. 333-366.
- IMF (2014): "Fiscal Policy and Income Inequality", *IMF Policy Paper*, 23, Washington DC.
- IZQUIERDO, J. y LACUESTA, A. (2012): "The contribution of changes in employment composition and relative returns to the evolution of wage inequality: the case of Spain", *Journal of Population Economics*, 25, pp. 511-543.
- IZQUIERDO, M. y LACUESTA, A. (2007): *Wage inequality in Spain: recent developments*, s.l.: Banco de España. Working Paper Series n°781.
- JIMENO, J., IZQUIERDO, M. y HERRANZ, V. (2001): "La desigualdad salarial en España: descomposición y variación por niveles de salarios", *Papeles de Economía Española*, Volumen 88, pp. 113-125.
- JONES, M., MARRON, J. y SHEATHER, S. (1996): "A brief survey of bandwidth selection for density estimation", *Journal of the American Statistical Association*, 91(433), pp. 401-407.
- KAKWANI, N. y PODDER, N. (2008): "Efficient estimation of the Lorenz curve and associated inequality measures from grouped observations", en: D. Chotikapanich, ed. *Modeling Income Distributions and Lorenz Curves*. s.l.:Springer, pp. 57-70.
- KATZ, L. y AUTOR, D. (1999): "Changes in the wage structure and earnings inequality", en: *Handbook of Labor Economics*. Amsterdam: North-Holland, pp. 1463-1555.
- KLEIBER, C. (2008): "A guide to the dagum distributions lorenz curve and associated inequality measures from grouped observations", en: D. Chotikapanich, ed. *Modeling Income Distributions and Lorenz Curves*. s.l.:Springer, pp. 97-111.
- KLEIBER, C. y KOTZ, S. (2003): *Statistical Size Distributions in Economics and Actuarial Sciences*. New York: Wiley.
- LEMIEUX, T. (2008): "The changing nature of wage inequality", *Journal of Population Economics*, pp. 21-48.

- LÓPEZ-LABORDA, J. y SANZ ARCEGA, E. (2012): “Las preferencias por la redistribución: Teoría y evidencia para España”, *Documentos de trabajo, FUNCAS*, p. 699.
- MERCADER-PRATS, M., PLANAS, M. y LEVY, H. (2001): “Una introducción a Espasim: un modelo de microsimulación para evaluar reformas del sistema de impuestos y subsidios sociales en España”, en: *Desigualdad, redistribución y bienestar: una aproximación a partir de la microsimulación de reformas fiscales*. Madrid: Instituto de Estudios Fiscales, pp. 39-60.
- MINOU, C. y REEDY, S. (2014): “Kernel density estimation on grouped data: the case of poverty assessment”, *Journal of Economic Inequality*, 12(2), pp. 163-189.
- MORAL-ARCE, I., DE LAS HERAS PÉREZ, A. y SPERLICH, S. (2019): “Recovering income distributions from aggregated data via micro-simulations”, *Spanish Journal of Statistics*, 1(1), pp. 13-29.
- MORAL-ARCE, I. y DE LAS HERAS, A. (2011): “An estimation of income distribution using grouped data: a general two-step”, *Papeles de Trabajo Instituto de Estudios Fiscales*, pp. 1-27.
- MUELLER, H., QUIMET, P. y SIMINTZI, E. (2017): “Within-firm pay inequality”, *The Review of Financial Studies* 30 (10), pp. 3605-3635.
- NADARAYA, E. (1964): “On estimating regression”, *Theory of Probability and its Applications*, 9(1), pp. 141-142.
- OECD (2014a): *Social and welfare issues: Inequality*. [En línea] Disponible en: <http://www.oecd.org/social/inequality.html>
- OECD (2014b): *Estudios económicos de la OCDE. España. Visión general*. Paris: OECD Publishing.
- OLIVER, X. y SPADARO, A. (2009): “GLADHISPANIA: un modelo de microsimulación del IRPF y de las cotizaciones sociales”, en: *Las reformas fiscales bajo el microscopio. Microsimulación fiscal en España: datos, metodología y aplicaciones*. Madrid: Instituto de Estudios Fiscales, pp. 159-178.
- ONRUBIA FERNÁNDEZ, J. y RODADO RUIZ, M. (2015): “El papel redistributivo del sistema fiscal: presente y futuro”, *Ekonomiaz*, 88 (2), pp. 176-217.
- ONRUBIA, J. y PICOS, F. (2011): “Comportamiento recaudatorio y redistributivo del IRPF en la última década”, *Economistas*, Volumen 126 (número extra "España 2010: un balance"), pp. 191-197.
- ONRUBIA, J. y PICOS, F. (2012): “Diseño, contenido y aplicaciones del nuevo panel de declarantes del IRPF 1999-2007”, *Revista de Economía Aplicada*, Volumen 60, pp. 53-87.
- ONRUBIA, J. y SANZ, J. (2003): Redistribución y bienestar a través de la imposición sobre la renta personal. *Estudios de Hacienda Pública. Instituto de Estudios Fiscales. Madrid*.
- PALACIO, J. y SIMÓN, H. (2004): Dispersión salarial entre establecimientos y desigualdad salarial en España. *Revista de Economía Aplicada*, XII(36), pp. 47-81.
- PARZEN, E. (1962): On estimation of a probability density function and mode. *The Annals of Mathematical Statistics*, 33(3), pp. 1065-1076.
- PERSSON, T. y TABELLINI, G. (1994): “Is inequality harmful for growth? Theory and evidence”, *American Economic Review*, 48, pp. 600-621.
- PICOS, F., DIAZ DE SARRALDE, S., MORENO, A. y TORREJÓN, L. (2009): “Microsim-IEF renta y la simulación de reformas fiscales: una aplicación a la reforma del IRPF 2007”, en: *Las reformas fiscales bajo el microscopio. Microsimulación fiscal en España: datos, metodología y aplicaciones*. Madrid: Instituto de Estudios Fiscales, pp. 139-158.

- PIJOAN, J. y SÁNCHEZ-MARCOS, V. (2010): “pain is different: falling trends of inequality”, *Review of Economics Dynamics* 13 (1), pp. 154-178.
- POTERBA, J. (2007): “Income inequality and income taxation”, *Journal of Policy Modelling*, Volumen 29, pp. 623-633.
- QUILIS, E. y FRUTOS, R. (2020): “Un análisis dinámico de la distribución salarial en España”, *Papeles de trabajo del Instituto de Estudios Fiscales* 3/20.
- REED, W. y WU, F. (2008): “New four- and five-parameter models for income distributions”, en: D. Chotikapanich, ed. *Modeling Income Distributions and Lorenz Curves*. s.l.:Springer, pp. 211-224.
- REYES, M., FRANCISCO-FERNANDEZ, M. y CAO, R. (2016): “Nonparametric kernel density estimation for generated grouped data”, *Journal of Nonparametric Statistics*, Volumen 28, pp. 235-249.
- REYES, M., FRANCISCO-FERNANDEZ, M. y CAO, R. (2017): “Bandwidth selection in kernel density estimation for interval-grouped data”, *TEST*, Volumen 26, pp. 527-545.
- REYES, M., FRANCISCO-FERNANDEZ, M. y CAO, R. (2019): “Kernel distribution estimation for grouped data”, *SORT*, 43(2), pp. 259-288.
- ROSENBLATT, M. (1956): “Remarks on some nonparametric estimates of a density function”, *The Annals of Mathematical Statistics*, 27(3), pp. 832-837.
- SÁEZ, E. (2001): “Using elasticities to derive optimal income tax rates”, *Review of Economics Studies*, Volumen 68, pp. 205-229.
- SÁEZ, E., SLEMROD, J. y GIERT, S. (2012): “The elasticity of Taxable Income with respect to marginal tax rates: a critical review”, *Journal of Economic Literature*, 50(1), pp. 3-50.
- SALA-I-MARTÍN, X. (2002): *The world distribution of income (estimated from individual country distribution)*, s.l.: National Bureau of Economic Research Working Paper, nº 8933.
- SALA-I-MARTÍN, X. (2006): “The world distribution of income: falling poverty and convergence period”, *Quarterly Journal of Economics*, 121(2), pp. 351-397.
- SANZ, J., ROMERO, D. y ÁLVAREZ, S. (2008): *La protección de la familia en la Unión Europea: análisis comparativo y simulación de reformas aplicadas a España*. s.l.: Fundación BBVA.
- SCOTT, D. y SHEATHER, S. (1985): “Kernel density estimation with binned data”, *Communications in Statistics, Theory and Methods*, Volumen 27, pp. 832-837.
- SERRANO-PUENTE, D. (2020): “Optimal progressivity of personal income tax: a general equilibrium evaluation for Spain”, *SERIEs*, Volumen 11, pp. 407-455.
- SILVERMAN, B. (1986): *Density estimation for statistics and data analysis*. s.l.:Chapman y Hall/CRC Monograph on Statistics and Applied Probability 26.
- SIMÓN, H. (2009): “La desigualdad salarial en España: Una perspectiva internacional y temporal”, *Investigaciones Económicas*, 33, pp. 439-471.
- SLEMROD, J. y BAKUJA, J. (2001): *Does growing inequality reduce tax progressivity? Should it?*, s.l.: National Bureau of Economic Research. Working Paper 7576.
- TITTERINGTON, D. (1983): “Kernel-based density estimation using censored, truncated or grouped data”, *Communications in Statistics, Theory and Methods*, Volumen 12, pp. 2151-2167.

- WALTER, P. y WEIMER, K. (2018): *Estimating poverty and inequality indicators using interval censored income data from the German microcensus*, Berling: Diskussionsbeiträge, No. 2018/10, Freie Universität Berlin, Fachbereich Wirtschaftswissenschaft.
- WAND, M. y JONES, M. (1995): *Kernel Smoothing*. s.l.:Monographs on statistics and Applied Probability. Springer-Science.
- WANG, B. y WERTELECKI, W. (2013): "Density estimation for data with rounding errors", *Computational Statistics & Data Analysis*, Volumen 65, pp. 4-12.
- WATSON, G. (1964): "Smooth regression analysis", *Sankhya: The Indian Journal of Statistics, Series A*, 26(4), pp. 359-372.
- WU, X. y PERLOFF, J. (2003): "Calculation of maximum entropy densities with application to income distribution", *Journal of Econometrics*, 115(2), pp. 347-354.
- XU, S. (2014): "Asymmetric kernel density estimation based on grouped data with applications to loss model", *Communications in Statistics-Simulation and Computation*, Volumen 43, pp. 657-672.