

Imposición sobre la renta personal: ¿qué sabemos de la relación entre redistribución y costes de eficiencia?*

Optimal personal income taxation: What do we know about the relationship between redistribution and efficiency costs?

Alejandro Esteller Moré

Universidad de Barcelona e Instituto de Economía de Barcelona

Resumen

La teoría de la imposición óptima ofrece recomendaciones sobre el diseño de un impuesto no-lineal sobre la renta que, minimizando los costes de eficiencia, maximice el bienestar dado un objetivo de redistribución. Para hacer operativas tales recomendaciones, es necesario disponer de la elasticidad de respuesta de los contribuyentes ante aumentos del tipo impositivo marginal, que junto con la forma de la distribución de la renta laboral y las preferencias sociales por la redistribución acaban determinando el diseño óptimo del impuesto. La teoría económica y los estimadores provistos en la literatura dotan, pues, al decisor social de un potente instrumental para guiar el diseño impositivo. Esto es lo que se revisa en este artículo de carácter panorámico.

Palabras clave: enfoque utilitarista, imposición óptima, elasticidad de la base impositiva.

Clasificación JEL: H21, H23, H24

Abstract

Optimal taxation theory offers recommendations on the design of a non-linear income tax that, while minimizing efficiency costs, maximizes social welfare given a redistribution target. To make such recommendations operational, it is necessary to know the taxpayers' response elasticity to increases in the marginal tax rate, which together with the shape of the labor income distribution and social preferences for redistribution end up determining the optimal design of the tax. Economic theory and the estimators provided in the literature, therefore, provide the social decision maker with powerful instruments to guide the tax design. This is what is reviewed in this panoramic article.

Keywords: utilitarian approach, optimal taxation, elasticity of taxable income.

* Quisiera agradecer los detallados comentarios y sugerencias de dos evaluadores, así como la financiación recibida del proyecto PID2021-126652NB-I00 del Ministerio de Ciencia e Innovación. En todo caso, cualquier error u omisión es responsabilidad exclusiva del autor.

1. Introducción

La desigualdad en la distribución de la renta, así como el análisis o simple consideración de las políticas públicas que pueden reducirla no fue un tema central para los economistas a lo largo del s. XX. En este sentido, por ejemplo, Sandmo afirmaba «this neglect is still visible in the allocation of space in introductory textbooks and books on microeconomic theory» (Sandmo, 2015, págs. 60-61), pues, ciertamente, en los manuales en pocas ocasiones se tiene en cuenta la heterogeneidad de los individuos en relación con su capacidad económica. En cambio, a lo largo de lo que llevamos de s. XXI, el trabajo de economistas tales como Atkinson, Stiglitz –con un amplio recorrido, en su caso, ya a lo largo del s. XX–, así como Piketty, Saez o Zucman ha puesto la redistribución en el centro del debate académico, pero también social. Por ejemplo, Christine Lagarde, cuando era directora gerente del FMI, afirmaba que «la investigación reciente del FMI nos dice que una menor desigualdad se asocia con una mayor estabilidad macroeconómica y mayor crecimiento económico sostenible» (Atkinson, 2015, pág. 12). Por tanto, las políticas de redistribución se han convertido también en un objetivo político de primer orden.

¿Es ese creciente interés acorde con un incremento en los niveles de desigualdad?. En la Figura 1, a título ilustrativo, mostramos la concentración de la renta (antes de impuestos, pero incluyendo las pensiones de jubilación) para el 1% más rico de la población para España, y para una serie de países desarrollados (EE. UU., Italia, Portugal y Francia) y para la Unión Europea (UE). En perspectiva histórica (hay datos para EE. UU. y para Francia desde la primera mitad del s. XX), se observa que la desigualdad decreció hasta, aproximadamente, 1980; desde entonces, la tendencia es creciente. Para los EE. UU., eso ha supuesto prácticamente volver a los niveles de desigualdad de inicio del s. XX. A diferencia del resto de países considerados, no se observa esa tendencia creciente en nuestro país (línea azul fuerte más gruesa), pero sí podemos comprobar cómo la desigualdad en España está ligeramente por encima de la de la UE, así como de la del resto de países incluidos en la Figura 1, excepto de la de los EE. UU. España es un país especialmente desigual en relación con la distribución de la renta.

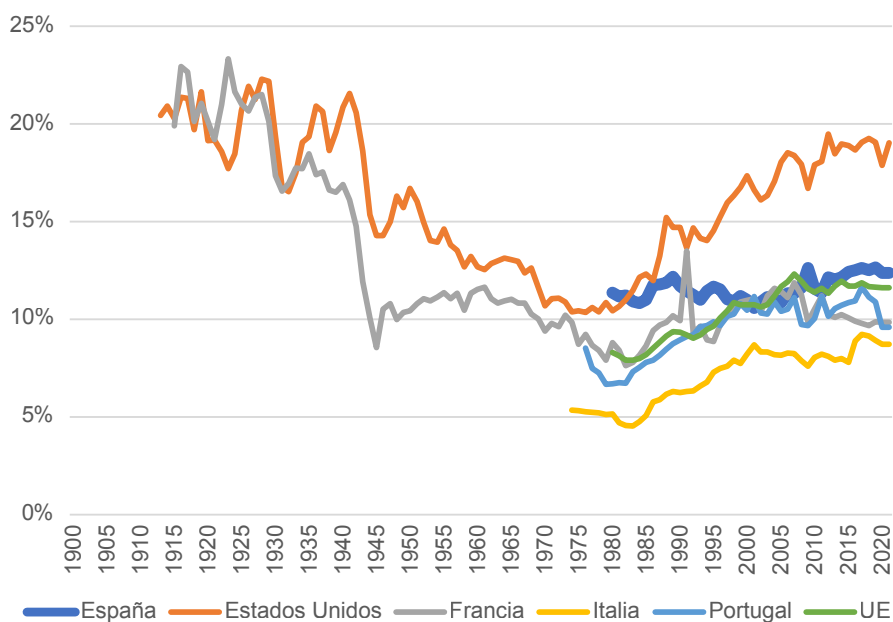
Más allá de los problemas que pueda haber por la dificultad intrínseca de trabajar con series largas y los supuestos que se han de realizar a la hora de construir estas bases de datos, la tendencia al aumento de la desigualdad ha llamado la atención de los economistas, generando un renovado interés en la profesión por los temas que tienen que ver con la desigualdad y con las políticas que la pueden reducir. Entre estas políticas, está el sistema impositivo y, en concreto, la imposición sobre la renta personal. Éste es el objeto de este artículo.

Gracias a Mirrlees (1971)¹, los economistas disponemos de un marco teórico para el diseño de un impuesto progresivo sobre la renta; como es de esperar, la potencia redistributiva del mismo depende en gran medida de cuál es la función objetivo del planificador social. Este modelo, a pesar de su valor indudable como

¹ Para una completa revisión técnica del modelo teórico propuesto por este autor, véase, por ejemplo, Salanié (2003), Capítulo 4.

marco pionero de análisis, no es de fácil aplicación (basado en el diseño de mecanismos para incentivar a los individuos a revelar su verdadera capacidad económica, siendo esta información privada); los resultados se basan en simulaciones numéricas, las cuales son muy dependientes de los supuestos realizados, que no son pocos (Tuomala y Weinzierl, 2020); y, en definitiva, sus prescripciones teóricas son difícilmente implementables, esto es, es complicado que sean socialmente aceptadas (básicamente, que el tipo marginal decrezca en los tramos elevados de la distribución o que éste apenas aumente) y no es fácil de caracterizar en términos de estadísticos suficientes (Diamond y Saez, 2011), concepto que definiremos más adelante. Las aproximaciones de Diamond (1998) y, sobre todo, de Saez (2001) intentan superar las problemáticas referidas. Este marco de análisis será el que consideraremos en este artículo para responder la pregunta planteada en su título revisando la literatura, básicamente, de carácter teórico.

FIGURA 1
CONCENTRACIÓN DE LA RENTA EN EL 1% MÁS RICO DE LA POBLACIÓN



FUENTE: World Inequality Database, acceso *on-line*: <https://wid.world/es/series/>

El resto del trabajo se estructura como sigue. En la sección 2, se lleva a cabo propiamente la revisión de la literatura teórica de la imposición óptima sobre la renta personal; después de realizar unas consideraciones previas (sección 2.1), se describe el modelo básico de Saez (2001) y se lleva a cabo una pequeña aplicación

para el caso español. En la sección 3, se revisan ciertas ampliaciones del modelo básico antes referido; y la sección 4 concluye.

2. Teoría de la imposición óptima sobre la renta personal

2.1. *Unas consideraciones previas*

Cualquiera que sea la aproximación a la teoría de la imposición óptima sobre la renta, ésta consiste en obtener los tipos marginales óptimos² a lo largo de la distribución de habilidades de los contribuyentes para obtener renta (en la práctica, aproximada la distribución de habilidades a partir de la observada para la renta salarial). A partir de esa distribución de tipos marginales, se pueden obtener los tipos medios que, para redistribuir, han de ser crecientes con esa capacidad³. Para ello, se parte de una función de bienestar social, la cual agrega las funciones individuales de utilidad de manera lineal (enfoque utilitarista) o bien dando un peso decreciente a los individuos con mayor capacidad económica, siendo la función rawlsiana de bienestar el caso límite, donde sólo el individuo con menor capacidad económica recibe un peso. En ausencia de información sobre la capacidad económica, concepto sobre el cual volveremos más adelante en esta misma sección, el gravamen de una «proxy» como es la obtención de renta genera, no obstante, costes de eficiencia. Y es aquí donde esos costes tienen que ser contrapuestos a las ganancias sociales de bienestar derivadas de la redistribución. Esta relación de intercambio quedará bien reflejada en el desarrollo matemático de la sección 2.2. y en el de la sección 2.3.

Ahora bien, antes de ese desarrollo, hay una cuestión previa que puede ayudar a entender lo que hay detrás de la teoría de la imposición óptima. Cuando aumenta el tipo marginal, se producen dos impactos: por un lado, aumenta el tipo medio de todos los individuos situados por encima del umbral a partir del cual aplica el marginal, y lo hace sin generar costes de eficiencia (no hay efecto sustitución, y se suele asumir que el efecto renta es nulo o despreciable) y aumentando el impacto redistributivo; y, por otro lado, ahora sí, genera un coste de eficiencia, el cual será más o menos importante dependiendo de cuántos individuos se localicen justo en el umbral (Jacobs, 2013)⁴. Por tanto, necesitamos saber cuántos individuos hay por encima de ese umbral (cuantos más haya y más intensa sea la preferencia social por

² En esta revisión, la función de tipos marginales que consideraremos será continua (el tipo varía –de manera creciente o decreciente– a lo largo de toda distribución de la renta), lo cual contrasta con las tarifas por escalones que observamos en la práctica. Véase, por ejemplo, Slemrod *et al.* (1994) para el diseño óptimo de una tarifa por escalones.

³ En estos modelos, hay una transferencia de suma fija para todos los individuos, que se financia con la recaudación del impuesto, y la cual tiende a desaparecer conforme aumenta la renta laboral; el tipo marginal implícito de esa transferencia puede ser, como veremos en la sección 3.1, elevado. Así, incluso ante un impuesto lineal, se produciría redistribución (i.e., el tipo medio efectivo crecería con la renta).

⁴ En las simulaciones numéricas para obtener el tipo marginal a lo largo de toda la distribución, éste es una función continua, en lugar de tramos, de manera que un aumento del tipo marginal en un punto concreto afecta a los individuos allí localizados.

la redistribución, mayor será el tipo marginal óptimo) y cuántos se localizan en el umbral (cuantos más haya y mayor sea la elasticidad de respuesta al tipo marginal, menor será el tipo marginal óptimo). Ése es el origen de la relación de intercambio, para cuya resolución se necesita saber cuál es la distribución de habilidades; este aspecto es del todo clave. Pero, además, como implícitamente se ha sugerido, se ha de conocer la respuesta de los contribuyentes ante un aumento del tipo marginal, lo cual determinará precisamente los costes de eficiencia del impuesto no-lineal.

Antes, hemos citado a James A. Mirrlees como el origen de la teoría de la imposición óptima, con una base matemática y enraizada en la teoría de la economía del bienestar. No obstante, podemos retrotraernos a John Stuart Mill, el cual, a mitad del s. XIX, promovió el enfoque utilitarista que, implícitamente, está detrás de la teoría de la imposición óptima (Boadway y Kuff, 2022). En concreto, de acuerdo con J. S. Mill, la carga tributaria que soporta cada contribuyente debe ser tal que cada uno realice el mismo sacrificio en términos de «medios» (*means*, en inglés en el original), lo cual debe ser entendido como igual pérdida de utilidad (Utz, 2002). Así, bajo la teoría del sacrificio marginal igual, el diseño de los impuestos sería equitativo (por definición, si la utilidad marginal de la renta es decreciente) y, además, maximizaría la eficiencia (i.e., minimizaría los costes de la imposición). Este enfoque contrasta, por ejemplo, con el igualitarismo, el cual promueve aumentos en la carga impositiva, aunque, en valor absoluto, las ganancias marginales de bienestar en favor del más desfavorecido sean inferiores a las pérdidas de bienestar del resto de la sociedad.

En contraposición con el enfoque utilitarista, existe el enfoque de la capacidad de pago y el del gasto personal. Bajo el primero, los impuestos a pagar por cada individuo han de depender de los recursos económicos que tengan a su disposición⁵. La medida estándar de esa capacidad es la denominada renta extensiva o comprehensiva o renta Schanz-Haig-Simons, por los autores que la promovieron (Schanz, 1896; Haig, 1921; Simons, 1938), y que se define como el nivel máximo de consumo anual que un individuo pudiera llevar a cabo sin ver disminuido su nivel inicial de riqueza. Este enfoque, que, por tanto, debiera incluir las herencias y donaciones recibidas, es el que aparece en los textos constitucionales que rigen el diseño de los impuestos en la práctica. La siguiente cuestión que plantea este enfoque es, entonces, cómo debe diferenciarse el gravamen entre individuos. Para ello, por un lado, tenemos el principio de equidad horizontal (idéntico nivel de imposición de los iguales) y el principio de equidad vertical (aquellos individuos con mayor capacidad deben tributar más). Para hacer operativos ambos principios, especialmente el segundo, es imprescindible caracterizar de alguna manera el bienestar de los individuos y suponer que las utilidades individuales se pueden comparar, de forma que el enfoque utilitarista y el de la capacidad económica tienden a coincidir a la hora de plantearse el diseño impositivo.

⁵ Hemos dicho «en contraposición», pues, por ejemplo, Henry Simons “*explicitly rejects justifications for redistribution that are based on differences in the value of income to the rich and the poor...*”, esto es, este autor rechazaba el diseño de los impuestos basándose en el enfoque utilitarista (citado por Pressman, 2018, pág. 157).

Bajo el enfoque del gasto anual personal, Kaldor (1955) consideró que los individuos deberían ser anualmente gravados por lo que utilizan para su consumo personal en lugar de lo que anualmente aportan en forma de renta. Curiosamente, Mill también estaba a favor de utilizar esta métrica para gravar a los individuos, al igual, por ejemplo, que el Informe Meade (1978), el cual proponía un impuesto anual progresivo sobre el gasto junto con un impuesto anual sobre la riqueza y otro sobre las herencias (Rubio Guerrero, 1986).

En la práctica, el impuesto con fines redistributivos más importante de que disponen los sistemas fiscales actuales es el impuesto sobre la renta personal –y éste, ya hemos dicho, es el objeto de este artículo–, cuyo diseño teórico se basa en modelos de imposición óptima donde los individuos son caracterizados por funciones individuales de utilidad (ahí, como hemos dicho, la conexión entre el enfoque basado en la capacidad de pago y el utilitarista) convenientemente agregadas en una función de bienestar social, y los cuales, claro está, se diferencian por su capacidad de pago. En otro caso, la progresividad y su consecuente redistribución perderían el sentido. A pesar de la preponderancia de un impuesto personal sobre la renta, en los sistemas fiscales, coexisten otros impuestos que acaban de complementar la capacidad económica gravada por el impuesto sobre la renta personal⁶. Así, por ejemplo, entienden Banks y Diamond (2010) que se ha de interpretar las recomendaciones que hizo el Informe Meade (1978).

2.2. Saez (2001): Teoría básica

Para entender el modelo de Saez (2001), partiremos de la situación más sencilla. Nos planteamos obtener el tipo marginal máximo que compatibilice los criterios de eficiencia (minimización del exceso de gravamen) y de equidad (redistribución, según el peso relativo dado en la función de bienestar social a los más desfavorecidos en términos de capacidad económica) ya referidos. Utilizaremos el análisis de perturbación tal que el óptimo es aquel en que los beneficios netos en términos de bienestar social de aumentar el tipo marginal en el tramo superior se agotan. A partir de ahí, en la sección 3.1, veremos que se puede generalizar esta formulación a lo largo de toda la distribución de habilidades de los contribuyentes.

Por un lado, cuando aumentamos infinitesimalmente el tipo marginal, dt , donde t es el tipo marginal sobre la renta, tenemos lo que se denomina un «efecto mecánico», M , esto es, un aumento de la recaudación impositiva:

$$M \equiv N (B_m - B_u) dt \geq 0 \quad [1]$$

⁶ De hecho, como veremos, en un marco de análisis intertemporal y en ausencia de herencias, un impuesto general sobre el consumo (con un tipo uniforme) se puede reformular como un impuesto proporcional sobre la renta que obtenga el individuo a lo largo de toda su vida. Sobre esto, siquiera de manera implícita, volveremos en la sección 3.2.

donde N es el número de individuos por encima del umbral superior, esto es, donde se sitúan los contribuyentes con mayor renta; B_m es la renta media de los contribuyentes situados en ese tramo, y B_u es el umbral mínimo de renta del tramo superior; por tanto, la renta gravada al nuevo marginal es $(B_m - B_u)$ y la recaudación extra es precisamente esa renta multiplicada por la variación del marginal y por el número de contribuyentes situados en el tramo superior. Éste es el beneficio marginal derivado de subir impuestos; en ausencia de cualquier otro efecto, el tipo marginal óptimo sería del 100%.

No obstante, esperamos que los contribuyentes reaccionen ante incrementos del tipo marginal; el «efecto de comportamiento», C , recoge esa circunstancia (suponiendo que no hay salto de tramo y que los individuos siguen trabajando):

$$C \equiv -N t (\partial B_m / \partial (1-t)) dt = -N (t/(1-t)) \varepsilon B_m dt \leq 0 \quad [2]$$

donde $\varepsilon \equiv (\partial B_m / \partial (1-t)) ((1-t)/B_m) \geq 0$ es la elasticidad de respuesta de la base (liquidable o la renta bruta; ya veremos en la siguiente sección) ante un aumento del tipo marginal neto, $1-t$, para los contribuyentes situados en el tramo superior; tal elasticidad es la media ponderada por el nivel de renta de las respuestas de ese grupo de contribuyentes⁷.

En definitiva, a partir del análisis marginal antes referido, tenemos que el tipo óptimo es aquél que, como dijimos, agota los beneficios netos en términos de bienestar derivados de aumentar t , esto es,

$$M + C = 0 = N (B_m - B_u) dt - N (t/(1-t)) \varepsilon B_m dt \quad [3]$$

Por tanto, el impacto marginal total sobre la recaudación, R , es inferior al impacto mecánico producido por la respuesta de los contribuyentes ante el aumento de t . A partir de esta constatación, podemos obtener el «exceso de gravamen» (EG) expresado como un porcentaje de M :

$$R \equiv M (1 + C/M) = M (1 - (t/(1-t)) \varepsilon a) = M (1 - EG(\%)) \leq M \quad [4]$$

donde $a \equiv B_m / (B_m - B_u)$. Saez (2001) demuestra, y éste es un aspecto novedoso y crucial de su aproximación, que la cola superior de la distribución de la renta salarial (como «proxy», recordemos, de las habilidades) se asemeja a una distribución de Pareto, tal que a es precisamente el «coeficiente de Pareto». Para este tipo

⁷ A lo largo del artículo, supondremos que el efecto renta es despreciable (Gruber y Saez, 2002), de forma que, implícitamente, la elasticidad que estamos utilizando es la elasticidad compensada, que, como veremos más adelante, es el estadístico suficiente para estimar los costes de eficiencia del impuesto (Chetty, 2009). No lo será cuando la respuesta de los contribuyentes implique externalidades fiscales tales como, por ejemplo, la transformación de rentas a gravar en el impuesto sobre la renta personal en rentas a gravar en el impuesto sobre sociedades, o cuando la elusión fiscal implique traspasar rentas gravables entre períodos impositivos. En esos casos, para obtener el coste de eficiencia de variar el tipo, se necesitará estimar otros estadísticos tal que la elasticidad compensada no será suficiente para estimar el exceso de gravamen o coste de eficiencia.

de distribución de habilidades, B_m/B_u tiende a una constante, de manera que a tampoco varía apenas a lo largo de B_u ⁸. Lo interesante es la interpretación de ese coeficiente, el cual, como hemos dicho, depende de la distribución de habilidades, que, como también ya sugerimos en la sección 2.1, es un aspecto clave. Un valor pequeño de a se corresponde con una cola de la distribución más gruesa, esto es, B_m/B_u tiende a ser mayor y, en consecuencia, hay una mayor concentración de la renta en el tramo superior. Dado un determinado valor de a , el valor de la elasticidad es un estadístico suficiente para obtener el $EG = (t/1-t)\varepsilon a \geq 0$, expresado como un porcentaje de M . Si $\varepsilon = 0$, el $EG = 0$ tal que el tipo óptimo tendería al 100%. Cuando a disminuye, por un lado, el efecto mecánico (M) aumenta, pues hay más renta en el tramo superior; por otro lado, el efecto de comportamiento (C) se mantiene invariado (efecto sustitución para los individuos localizados en ese tramo). Por tanto, el efecto neto es tal que el EG , definido como C/M , disminuye; de ahí, el impacto positivo de a sobre el EG ⁹. En definitiva, el exceso de gravamen aumenta con el valor de la elasticidad (compensada), el valor del tipo marginal óptimo y el coeficiente Pareto, a .

En relación con el «efecto mecánico», obsérvese que se trata de una transferencia de renta desde manos de los contribuyentes situados en el tramo superior al sector público para redistribuir. Implícitamente, hasta el momento, hemos estado suponiendo que «todo» es ganancia para el sector público o, dicho de otra manera, la utilidad marginal de la renta en manos del sector público es infinitamente superior a la que el planificador social otorga a una unidad adicional de renta en manos de ese grupo de contribuyentes. Si suponemos que eso no es así, no todo es ganancia, y hemos de ponderar el componente M en la expresión [3] tal que, ya habiendo simplificado algunos de sus componentes, pasamos a tener:

$$M' + C = 0 = (1-p) (B_m - B_u) - (t/1-t) \varepsilon B_m \quad [5]$$

donde p es la ponderación relativa de la utilidad marginal social de la renta en manos del sector público versus si ésta está en manos de los contribuyentes de mayor capacidad económica¹⁰. Como es habitual, para que el tipo marginal

⁸ El umbral a partir del cual se sitúa el tramo superior, siguiendo el análisis coste-beneficio planteado en 2.1, tampoco debería ser muy elevado. Situar el marginal máximo en niveles muy elevados sólo hace aumentar los costes de eficiencia entre esos individuos –probablemente los que más responden– y reducir la capacidad redistributiva del sistema. Por ello, Jacobs (2013) propone que el tramo superior empiece, aproximadamente, a partir de la moda de la renta. Sin duda, ello es del todo «razonable» si lo que se quiere es tener bases amplias para redistribuir, pero, probablemente, no sea fácilmente «aceptable» socialmente. Éste, pienso, es un aspecto especialmente importante para el diseño del impuesto en el caso español, donde las reformas se focalizan en incrementar el tipo marginal máximo a partir de niveles muy elevados de base liquidable.

⁹ Véase, por ejemplo, Slemrod y Gillitzer (2014), Capítulo 3, Sección 5.1.

¹⁰ Obsérvese que el coste marginal de los fondos públicos (CMgFP) es $(M/(M+C))$ y, en equilibrio, éste se ha de igualar con el beneficio marginal, $1/p$, dada la definición de p , que es la ratio entre la utilidad marginal de la renta para el grupo de contribuyentes en el tramo superior y la utilidad marginal de la renta para el planificador social. Si se iguala $BMgFP = CMgFP$, tenemos $(1-p) \times M + C = 0$, que es exactamente la expresión [5], de donde podremos obtener el tipo marginal óptimo. Recordemos que $C \leq 0$ y $p < 1$.

sea estrictamente positivo, es necesario que $p < 1$. En [3], recordemos que $p = 0$, siendo ésta la situación conocida en la literatura como «soaking the rich» (acorde con unas preferencias sociales rawlsianas); el planificador social no valora en absoluto la pérdida de renta de los más ricos, de manera que sólo busca maximizar la recaudación obtenida de ese grupo de contribuyentes para redistribuir hacia el resto de la población (esto es, los relativamente más desfavorecidos). Estrictamente, la maximización para $p = 0$ nos dará el máximo de la conocida «curva de Laffer».

A partir de [5], ahora sí, podemos obtener el tipo marginal óptimo para los contribuyentes situados en el tramo superior (a lo largo de éste, el impuesto sí es lineal), t^* :

$$t^* = (1-p) / (1-p + (a \varepsilon)) \quad [6]$$

Dado un determinado valor de ε , cuanto menor sea el valor de a , *ceteris paribus*, mayores son los beneficios de aumentar el tipo marginal de la renta en aras de la redistribución personal.

La expresión [6] es, por tanto, muy intuitiva y nos ofrece unas recomendaciones claras sobre los factores de los cuales depende el tipo marginal óptimo máximo sobre la renta. Éste será tanto mayor, cuanto:

- (i) Menor sea el coeficiente de Pareto (esto es, más renta se concentre por encima de B_u);
- (ii) Menor sea el valor social de la renta en manos de los individuos ricos respecto de que ésta esté en manos del sector público (caso límite: $p=0$); y
- (iii) Menor sea la elasticidad de respuesta de los contribuyentes ante un incremento del tipo marginal neto.

Podemos obtener unos resultados particulares. En el caso límite en que $p = 1$, el tipo marginal máximo es nulo; y lo mismo sucede en el caso en que la elasticidad tienda a infinito. Para el individuo situado en el extremo de la distribución, $B_m - B_u$ tiende a cero, de forma que $M = 0$, y al haber sólo un efecto negativo sobre el comportamiento (no hay nadie por encima de B_u tal que pueda contribuir a la redistribución, como explicamos en 2.1), el tipo marginal óptimo únicamente para ese individuo es nulo (Mirrlees, 1971)¹¹. Y, en el caso de un impuesto lineal sobre la renta, $B_u = 0$, [6] se simplifica tal que acabamos obteniendo una regla «tipo Ramsey», esto es, $t^* = (1-p) / (1-p + \varepsilon)$.

Volviendo a [6], necesitamos, por tanto, conocer el valor de ε , dados los valores de p y de a . En la medida en que el planificador sólo tiene poder o capaci-

¹¹ Este resultado es conocido como «no distorsión en la cima» de la distribución, y, si bien es contra-intuitivo y de poca aplicabilidad práctica, es fácilmente explicable por el análisis coste-beneficio marginal que acabamos de comentar.

dad para variar el tipo marginal, da igual cuál sea el origen de las respuestas del contribuyente ante un aumento del tipo, de manera que, en general, la elasticidad compensada se convierte en el estadístico suficiente para conocer el exceso de gravamen (ver nota al pie 7) (Chetty, 2009). A la elasticidad que incorpora las respuestas debidas al comportamiento, a la evasión o a la elusión fiscal¹² se le denomina «elasticidad de la base liquidable» (en inglés, elasticity of taxable income, ETI)¹³.

2.3. Saez (2001): Ejemplo práctico básico

Afortunadamente, la literatura empírica ha avanzado mucho en los últimos tiempos y, en la actualidad, se dispone de estimaciones robustas de la ETI, ε (Neisser, 2021). Ahora bien, contra la validez interna de los estimadores, se ha de contraponer su validez externa. La elasticidad estimada es contingente a los parámetros legales del impuesto de cada país incluyendo los esfuerzos de la administración tributaria en controlar las posibilidades de fraude o de elusión fiscal (Slemrod y Kopczuk, 2002) o, simplemente, a la moral impositiva de cada país.

De acuerdo con el reciente metanálisis de Neisser (2021) (diferentes países, diferentes metodologías y diferentes periodos)¹⁴, la elasticidad de respuesta al tipo marginal neto (después de reducciones de la base) se sitúa en promedio en el 0,403, pero con un rango de valores que va desde el 0,074 al 0,887. Aparte de dudar de la validez externa para el caso español de esos estimadores, no se proveen elasticidades según la naturaleza de la renta. Esa información sí la tenemos para el caso español; Almunia y López-Rodríguez (2019) estimaron una elasticidad de la base liquidable de entre 0,18-0,38 para la renta laboral; mucho menor para la renta bruta declarada (suma de ingresos íntegros, antes de la aplicación de gastos fiscalmente deducibles o de reducciones), que se sitúa entre el 0,05 y el 0,08. Dado este rango de valores, vamos a realizar un sencillo ejercicio numérico utilizando cuatro valores: 0,08, 0,15, 0,25 y 0,4.

¹² La diferencia entre evasión y elusión fiscal no es muy clara en todas las ocasiones. En este artículo, siguiendo a Slemrod y Gillitzer (2014), Sección 3.3, consideramos que ese concepto incluye desde el arbitraje impositivo (entre tipos de renta, entre periodos impositivos o entre formas de organización de la actividad económica) hasta la planificación fiscal.

¹³ Un aspecto que no revisaremos es cómo la respuesta en términos de movilidad internacional puede afectar al diseño de los impuestos. En estos casos, será necesario, por tanto, estimar también esta elasticidad de respuesta de naturaleza extensiva. Véase Piketty y Saez (2013) para la extensión del modelo básico descrito; y Kleven *et al.* (2020) para, bajo este marco, el diseño de regímenes especiales que atraigan inmigrantes que provoquen externalidades positivas en el país de acogida.

¹⁴ La naturaleza de las respuestas es intensiva, que es el objeto principal de este artículo. Véase, por ejemplo, Blundell *et al.* (2011) para una estimación de elasticidades extensivas; y la reciente revisión de la literatura de Kleven (2021) focalizada en los incentivos a permanecer en el mercado laboral relacionados con impuestos negativos sobre la renta laboral en los tramos inferiores de la distribución.

La elasticidad de la renta bruta declarada recoge los efectos del impuesto sobre la evasión, así como efectos reales de comportamiento (*i.e.*, reducción del número de horas trabajadas), mientras que la de la base liquidable (ETI) recoge, además, el impacto de la elusión fiscal (de ahí, su mayor valor; recordemos que son elasticidades positivas, pues están expresadas en relación con el tipo marginal neto, $1-t$)¹⁵. A priori, la elusión fiscal se puede llegar a reconducir a través de un «mejor» diseño del impuesto, o ésta debería ser la política en lugar de reducir tipos (Diamond y Saez, 2011)¹⁶. El valor de la elasticidad de 0,08 nos servirá, pues, de punto de referencia para saber cuál podría llegar a ser el tipo óptimo si se eliminaran las posibilidades de elusión fiscal; eso sí, en la práctica, tal acción debe hacerse sin descartar que posibles beneficios fiscales puedan estar legitimados y, por tanto, la posibilidad de eliminar algunos de esos beneficios fiscales está limitada (Saez *et al.*, 2012). Por último, utilizaremos diferentes valoraciones sociales relativas de la renta en posesión de los ricos, $p = 0$ (que, recordemos, nos dará el máximo de la conocida «curva de Laffer», donde el planificador sólo se preocupa de maximizar la recaudación de los individuos situados en el tramo superior) y $p = 0,25$ (una unidad adicional de renta en manos del sector público para redistribuir tiene un valor social cuatro veces superior a si esa unidad de renta está en manos de los contribuyentes más ricos). Utilizaremos un coeficiente de Pareto igual a 3, de acuerdo con los cálculos de Esteller (2021), pero también, a modo ilustrativo, utilizaremos variación territorial en ese coeficiente. Los resultados se muestran en la Tabla 1. En rojo, se muestran las situaciones donde el tipo marginal real (calculado teniendo en cuenta la presión fiscal vía IRPF e IVA, el cual acaba también gravando la renta) se sitúa por encima del óptimo.

¹⁵ Véase, por ejemplo, Harju y Matikka (2016).

¹⁶ Por tanto, para reducir los efectos negativos de la elusión fiscal, la respuesta óptima es la bajada de tipos sólo si los beneficios fiscales están bien justificados en base a motivos sociales o, en general, en el de incentivar ciertos comportamientos tal que no es conveniente eliminarlos. Esta posible reducción de tipos contrasta con la situación planteada por Piketty *et al.* (2014). En algunas profesiones de alta cualificación, los salarios se fijan en base a una negociación vis a vis empleador-empleado tal que éste último consigue obtener unas rentas extraordinarias (*i.e.*, se le paga por encima de su productividad marginal). Para evitar los incentivos a la negociación, en estos casos, lo razonable es precisamente subir los tipos impositivos. Ello hace que el tipo marginal óptimo en el tramo superior debiera ser incrementado en función del valor de la elasticidad que recoge el aumento de la compensación relacionado con los esfuerzos bilaterales de negociación (su estimación empírica no parece una cuestión menor). Así, las menores rentas percibidas por esos trabajadores se distribuyen –se supone– entre el resto de la población y, además, desaparecen los esfuerzos socialmente improductivos de «búsqueda de rentas». Estos son los motivos que, en términos de bienestar social, hay detrás del incremento de tipos en el óptimo.

TABLA 1
TIPOS MARGINALES MÁXIMOS ÓPTIMOS

	Total Nacional ($\alpha=3$)			
<i>Tipo real</i> = 50,7%	$\varepsilon = 0,08$	$\varepsilon = 0,15$	$\varepsilon = 0,25$	$\varepsilon = 0,4$
$p=0$	80,6%	69,0%	57,1%	45,5%
$p=0,25$	75,18%	62,5%	50,0%	38,5%
	C. A. de Madrid ($\alpha=2,53$)			
<i>Tipo real</i> = 54,5%	$\varepsilon = 0,08$	$\varepsilon = 0,15$	$\varepsilon = 0,25$	$\varepsilon = 0,4$
$p=0$	83,2%	72,5%	61,3%	49,7%
$p=0,25$	78,7%	66,4%	54,2%	42,6%
	Navarra ($\alpha=4,2$)			
<i>Tipo real</i> = 60,3%	$\varepsilon = 0,08$	$\varepsilon = 0,15$	$\varepsilon = 0,25$	$\varepsilon = 0,4$
$p=0$	74,9%	61,3%	48,8%	37,3%
$p=0,25$	69,1%	54,3%	41,7%	30,9%

NOTA: el tipo (marginal) real es la «suma» del gravamen marginal máximo de IRPF y del tipo general del IVA, que implícitamente grava la renta (de hecho, se aplica sobre la renta disponible). Para la definición de ese tipo real, no hemos considerado las cotizaciones a la Seguridad Social, pues suponemos que, para los individuos situados en el tramo superior, aplica la cuota máxima de cotización. Así, el tipo real es ($Mg\ IRPF + \text{Tipo general IVA}$)/(1+ Tipo General IVA), siendo Tipo general IVA=21% para todo el territorio nacional (véase, por ejemplo, Jacobs, 2013, nota a la Tabla 1; Kleven *et al.*, 2020, nota al pie 4). El $Mg\ IRPF$ varía entre territorios; las fuentes de éste se dan seguidamente. En rojo, se identifican aquellas situaciones donde el tipo socialmente óptimo se sitúa por debajo del real.

FUENTES: para el Total Nacional (suponiendo que las CC. AA. no han actuado sobre la tarifa autonómica), el marginal máximo de IRPF es la suma del estatal (<https://sede.agenciatributaria.gob.es/Sede/ayuda/manuales-videos-folletos/manuales-practicos/irpf-2021/capitulo-15-calculo-impuesto-determinacion-integras/gravamen-base-liquidable-general/gravamen-estatal.html>) más el autonómico (en ausencia de acción legislativa por la C. A., se aplica el 15,87% según el art. 74.1.1ª de la Ley 35/2006 del IRPF), esto es, 40,37%; para la C. A. de Madrid, el marginal máximo es del 20,5% (autonómico) más el 24,5% (estatal), <https://www.comunidad.madrid/servicios/atencion-contribuyente/impuesto-renta-personas-fisicas>; y para Navarra, tenemos que el tipo marginal máximo foral es del 52%, <https://www.navarra.es/NR/rdonlyres/BFFF5E53-146A-4CAC-9A6B-F9EF3C00E8F9/0/tarifasIRPF.pdf>

Para el total nacional y una elasticidad elevada ($\varepsilon = 0,4$), observamos que el tipo óptimo de la «curva de Laffer» ($p = 0$) se sitúa en el 45,5%, por debajo del máximo real, 50,7%; mientras que, si sólo se considerara el tipo real marginal máximo de IRPF, 40,37%, quedaría por encima. Esto es, considerando el impacto del IVA, suponiendo una elasticidad relativamente elevada ($\varepsilon = 0,4$) e independientemente del peso dado a los ricos (p entre 0 y 0,25), concluimos que estamos en la parte descendente de la «curva de Laffer». Estos resultados, en cambio, no se dan para elasticidades no tan elevadas; de ahí, la importancia de la estimación de ese estadístico. Por ejemplo, si nos vamos a un valor de la elasticidad tal que $\varepsilon = 0,15$, independientemente del peso dado a los contribuyentes más ricos, el tipo óptimo socialmente se encuentra muy por encima del real; por ejemplo, el máximo de la «curva de Laffer» es del 69,0%. Y si se eliminan, bajo las consideraciones antes referidas, las posibilidades de elusión fiscal ($\varepsilon = 0,08$; sólo hay efectos reales sobre el comportamiento y la eva-

sión), el máximo de la «curva de Laffer» pasa a ser del 80,6%, muy por encima del real, del 50,7%. Estos resultados son en ausencia de actividad legislativa sobre la tarifa general por parte de las CC. AA. de régimen común.

Dado que la definición de la base imponible y de la base liquidable no difieren entre CC. AA. de régimen común y los criterios y esfuerzos de la administración tributaria debieran ser homogéneos entre esas CC. AA., no es de esperar grandes diferencias entre territorios en las elasticidades de respuesta; y, en cualquier caso, no consideramos esa posibilidad. En cambio, sí se observan discrepancias en la distribución de habilidades aproximadas por la renta salarial (Esteller, 2021); en este ejercicio, esas diferencias se recogen en el coeficiente de Pareto. Por ello, en la Tabla 1, hemos considerado las diferencias que pueda haber en los tipos óptimos por CC. AA. Y éstas existen, y no son desdeñables. Así, por ejemplo, para $\varepsilon = 0,15$, el máximo de la «curva de Laffer» para Navarra es el 61,3%, y para la C. A. de Madrid del 72,5%, mientras que el óptimo para todo el territorio nacional (incluyendo también Navarra) se sitúa en el 69,0%¹⁷. Por tanto, según esta teoría, la descentralización fiscal –dejando de lado los efectos nocivos que pueda provocar la competencia fiscal– debiera permitir, al modo del «teorema de la descentralización» por el lado del gasto (Oates, 1999), realizar las ventajas, en este caso, de la diferenciación impositiva entre territorios.

3. Extensiones

3.1. Tipos marginales a lo largo de toda la distribución de la renta

El marco teórico básico de la sección 2.2. se puede extender a lo largo de toda la distribución de contribuyentes. Así, se puede comprobar si los tipos marginales crecen a lo largo de toda la distribución, o tienen forma de U o de U-invertida. De hecho, demostrar que, bajo supuestos razonables e independientemente de las preferencias sociales por la redistribución, la forma funcional es de U fue el valor añadido de Diamond (1998), cuyo modelo fue posteriormente desarrollado y hecho, digamos, tratable por Saez (2001). A diferencia de lo que sugerían las simulaciones iniciales de Mirrlees (1971), una distribución con forma de U-invertida o un impuesto lineal junto con una transferencia de suma fija universal no es óptimo.

Por tanto, según esta nueva línea de investigación, los tipos marginales óptimos son elevados para los tramos inferiores de la distribución, disminuyen para los tramos medios y vuelven a aumentar para los tramos elevados. Como vamos a ver, de nuevo, esto tiene que ver con lo que comentábamos en la sección 2.1, la densidad

¹⁷ Si, para $\varepsilon = 0,15$, calculamos el EG (como porcentaje del efecto mecánico), $EG(\%) \equiv (t/1-t)\varepsilon a$ (véase expresión [4]), obtenemos que éste es un 46,3%, 45,5% y 95,7% para todo el territorio nacional, la C. A. de Madrid y Navarra, respectivamente. Un $EG > 100\%$ indica que estamos en la parte descendente de la «curva de Laffer»; mientras que, cuanto más bajo sea ese porcentaje, más margen hay para subir los impuestos.

de la renta por tramos, así como la concentración de ésta por encima de cada punto de la distribución. Esto es, la forma de la distribución de habilidades es fundamental, en particular, para la cola superior (distribución de Pareto).

Diamond (1998) obtiene, suponiendo que la función de utilidad de los individuos es cuasilineal y así eliminando, de nuevo, el «efecto renta», la siguiente fórmula de tipos marginales para todos los niveles de renta:

$$t(B)/(1-t(B)) = (1/\varepsilon(B)) RR (1-P(B)) \quad [7]$$

A diferencia de la situación analizada en 2.2, el tipo marginal, t , depende aquí de la renta, B ; allí, era el mismo para todos los contribuyentes situados en el tramo superior. Definimos RR como la ratio de riesgo (en inglés, *hazard ratio*), tal que $RR \equiv (1-H(B))/(B h(B))$, donde $H(B)$ es la distribución acumulada de contribuyentes hasta B y $h(B)$ es la densidad de éstos alrededor de B . Por tanto, RR recoge de manera explícita el coste-beneficio marginal de variar el tipo marginal ya identificado en la sección 2.1. *Ceteris paribus*, cuanto mayor es la concentración de contribuyentes por encima de B , $1-H(B)$, y menor es la densidad de la renta alrededor de B , $B h(B)$, mayor debe ser el tipo marginal (la parte izquierda de [7]). De nuevo, por tanto, observamos cuán importante es la forma funcional de la distribución de la renta laboral. A partir de ahí, de nuevo, se observa claramente que el tipo óptimo es menor cuanto mayor sea la elasticidad de respuesta de los individuos con una renta igual a B . Y, por último, $1-P(B)$ identifica el valor social promedio de la renta para los contribuyentes con una renta inferior a B ; por tanto, recoge las preferencias sociales por la redistribución¹⁸. Cuanto más elevado, mayor será el tipo marginal en el óptimo. Boadway y Kuff (2022) reexpresan [7] para que quede más clara la relación de intercambio entre eficiencia y equidad:

$$t(B)/(1-t(B)) = [(1-H(B))(1-P(B))]/[\varepsilon(B) B h(B)] = \text{equidad} / \text{eficiencia} \quad [7']$$

A partir de la distribución salarial, RR tiene forma de U ¹⁹. Por tanto, *ceteris paribus*, la forma de la distribución de habilidades genera esa forma funcional en forma de U para la distribución de los tipos marginales. Para niveles elevados de renta, la distribución es de Pareto y a coincide con RR , pero los tipos marginales siguen aumentando por el menor peso dado a esos individuos en la función de bienestar social²⁰.

¹⁸ No es óptimo que t –en aras de redistribuir– sea negativo; por tanto, siempre será más eficiente imponer $t > 0$ y redistribuir mediante transferencias de suma fija con la recaudación obtenida. Para demostrarlo, volvamos a la técnica de la perturbación partiendo de $t < 0$. Si aumentamos marginalmente t , el efecto mecánico es positivo, en tanto en cuanto $P(B) < 1$, mientras que el efecto sobre el comportamiento, en este caso, será positivo, ya que $1-t < 0$. Ambos efectos van en la misma dirección de incrementar el tipo partiendo éste de un valor negativo. Es óptimo, pues, que $t > 0$.

¹⁹ Ver, por ejemplo, Gráfico 2 de Diamond y Saez (2011). De hecho, grafican $1/RR$; de ahí, la forma de U invertida en ese gráfico.

²⁰ Podría ser que, a diferencia del supuesto que se suele emplear en las simulaciones numéricas, la elasticidad de respuesta fuese mayor entre los individuos situados en la cola superior de la distribución. Ello haría que los

Los elevados tipos marginales en la cola inferior permiten incrementar el importe de la transferencia de suma fija hacia esos individuos (véase nota al pie 18) al aumentar el tipo medio que soportan los tramos superiores sin generar desincentivos a esos individuos a trabajar y donde los costes de eficiencia para los relativamente pobres son bajos precisamente por el pequeño importe de sus bases. Eso plantea una peculiar relación de intercambio en la cola inferior: en general, no es posible reducir la «trampa de la pobreza» (los elevados tipos marginales generan desincentivos a trabajar para los más pobres) sin sacrificar la reducción de la pobreza (Jacobs, 2013). Éste es un aspecto importante que tener presente al diseñar la tarifa en los tramos inferiores de la distribución: paradójicamente, cuanto mayor sea el deseo de redistribución, mayores deberán ser los tipos marginales para los contribuyentes de la cola inferior de la distribución.

Partiendo de estos niveles elevados del tipo marginal en el tramo inferior, de nuevo, la tendencia decreciente hasta la moda sigue la resolución de la relación de intercambio planteada por la evolución de la *RR*: mayor densidad (mayores costes de eficiencia) vs. menor acumulación de contribuyentes por encima de esos valores (menores beneficios de la redistribución). Esto es así hasta llegar a la moda a partir de la cual, de nuevo, por los motivos antes descritos (menor densidad y menor peso en la función de bienestar social), los tipos marginales vuelven a aumentar hasta llegar al contribuyente, si es posible identificarlo, de mayor habilidad, cuyo tipo marginal óptimo es nulo.

3.2. Imposición óptima sobre los rendimientos del ahorro

Hasta el momento, hemos estado considerando sólo el gravamen óptimo sobre la renta laboral. No obstante, *a priori*, la capacidad económica debe considerar también la renta derivada del ahorro (sección 2.1). Suponiendo que el precio de mercado de los bienes o servicios hoy y mañana es el mismo, el coste de oportunidad de renunciar a una unidad de consumo hoy para consumir en el futuro es $1/(1+r)$, donde r es la tasa de descuento, que por ahora suponemos es idéntica al tipo de interés de mercado. Un impuesto sobre los rendimientos del ahorro hace que esa rentabilidad pase a ser $r(1-t)$, de forma que, implícitamente, el gravamen del ahorro se puede asimilar a un impuesto sobre el consumo futuro, $1/(1+r) < 1/(1+r(1-t))$.

Entendido esto, cabe, por tanto, plantearse si es óptimo gravar de manera distinta el consumo futuro y presente; si la respuesta es afirmativa, a pesar de la doble imposición que supone un gravamen del ahorro²¹ (pues la renta que generó la capacidad de ahorrar ya fue gravada, y la rentabilidad obtenida, por ahora, es simple-

marginales óptimos fuesen menores para ese grupo de contribuyentes. No obstante, el supuesto de una misma elasticidad se justifica por considerar que la mayor respuesta en los tramos elevados de la distribución se debe, principalmente, a comportamientos de elusión fiscal (Diamond y Saez, 2011, pág. 174).

²¹ Entre otros motivos, esto puede justificar la existencia de sistemas fiscales duales (véase, por ejemplo, Sørensen, 2005).

mente la compensación por diferir el momento del consumo) y de la distorsión extra provocada, debe haber algún argumento relacionado con la capacidad redistributiva del sistema fiscal. Para ello, analizaremos, de entrada, la conveniencia de gravar el consumo en general, para volver luego al análisis del gravamen de los rendimientos del ahorro, que es lo que nos ocupa en esta sección.

Al respecto, lo primero que hay que entender, e implícitamente ya lo hemos utilizado en la Tabla 1, es que un impuesto uniforme (i.e., un tipo único fijo) sobre el consumo es equivalente a un impuesto sobre la renta laboral. Por tanto, si la uniformidad en el gravamen del consumo es óptima, ese impuesto no añade nada en términos de bienestar social tal que se puede conseguir exactamente lo mismo aumentando equiproporcionalmente los tipos impositivos sobre la renta a lo largo de toda la distribución de contribuyentes. La situación paradigmática en que sí es socialmente óptimo un impuesto sobre el consumo se da cuando las preferencias entre ocio y la composición de la cesta de consumo no son débilmente separables (Deaton, 1979), esto es, el número de horas dedicadas al ocio está correlacionado con la composición (no el nivel) de la cesta de consumo.

La intuición que hay detrás de este resultado es que, en un mundo de información imperfecta tal y como el utilizado por Mirrlees (1971), a quien queremos gravar más fuertemente es a los individuos de elevada productividad. Éstos, no obstante, ante impuestos elevados, decidirán trabajar menos y obtener menos renta haciéndose pasar por individuos de baja productividad. Tenderán a consumir demasiado ocio, pues pueden hacerse pasar por un individuo de baja productividad (obtener igual renta) trabajando menos que aquéllos. Para evitar ese incentivo, que va en contra del objetivo de redistribución, el planificador social puede gravar el ocio (que, por lo explicado antes, resulta una señal de productividad); dado que eso no es posible, si la composición de la cesta de consumo es dependiente del nivel ocio, es inmediato inferir que un impuesto general sobre el consumo que grave más intensamente los bienes y servicios complementarios con el ocio es óptimo (Atkinson y Stiglitz, 1976). La combinación de un impuesto no lineal sobre la renta y un impuesto sobre el consumo aumenta la redistribución y, por tanto, el bienestar social bajo ese esquema peculiar de las preferencias individuales²².

De ahí, es inmediata la analogía con el gravamen de los rendimientos del ahorro. En concreto, si la tasa de descuento de los individuos más hábiles es menor y, por tanto, muestran una mayor propensión al ahorro, de nuevo, el gravamen de los rendimientos del ahorro es una manera implícita de gravar a los individuos de elevada productividad (e.g., Saez, 2002) y, por tanto, de elevar los niveles de redistribución. Ciertamente, existen otros argumentos para gravar los rendimientos normales del ahorro (Jacobs, 2013, Sección 5), como pueden ser las dificultades existentes a la hora de diferenciar rentas del trabajo y del capital (Piketty y Saez, 2013,

²² Ciertamente, en la práctica existen otras imperfecciones más allá de la información asimétrica. Por ejemplo, en presencia de fraude en un impuesto no-lineal sobre la renta, independientemente de la estructura de preferencias individuales, puede ser óptima la existencia de un impuesto general sobre el consumo (Boadway *et al.*, 1994).

sección 3.4); y, en cualquier caso, desde el punto de vista de la eficiencia impositiva, siempre es deseable, si las hay, gravar las rentas económicas, esto es, los rendimientos que superan los rendimientos normales (o tasa de descuento). En este último caso, no habrá una relación de intercambio entre eficiencia y equidad.

En definitiva, repetimos, a pesar del doble gravamen que supone gravar la renta que permite ahorrar y el posterior rendimiento del ahorro, la redistribución y el bienestar social mejorarán si a mayor productividad, mayor propensión al ahorro. Esto es importante para justificar el gravamen de los rendimientos del ahorro, cuyo nivel dependerá de otros factores como puede ser la movilidad del capital. En la actualidad, disponemos de marcos teóricos que nos permiten calcular, de manera similar a lo realizado en la sección 2.2., los tipos óptimos sobre los rendimientos del capital (Saez y Stantcheva, 2018). En el caso español, a la hora de realizar ese ejercicio, debe tenerse en cuenta, en cualquier caso, la presión fiscal que el capital soporta o puede soportar en el impuesto sobre el patrimonio neto²³.

4. Conclusiones

En este artículo, hemos revisado la literatura de la imposición óptima, en concreto, en relación con un impuesto sobre la renta personal. A diferencia de la literatura iniciada por Mirrlees (1971), las aproximaciones de finales del siglo pasado e inicios del actual combinan de una manera muy intuitiva y clara la teoría con su aplicación práctica. Ello dota de un instrumental muy potente a los decisores sociales a la hora de diseñar este impuesto, eso sí, resolviendo la relación de intercambio eficiencia-equidad. Este *trade-off* sigue existiendo; quién sabe si en el futuro dejará de existir, de forma que los impuestos pasen a ser de suma fija, esto es, basados en la capacidad económica, y no en una aproximación de ésta, como es la renta.

La revisión de la literatura que hemos realizado no ha sido exhaustiva, sino que hemos intentado identificar los puntos más importantes, de forma que el lector pueda entender los mecanismos básicos que hay detrás de estos modelos de imposición óptima. En este sentido, la revisión se ha centrado en los aspectos teóricos, aunque, por lo dicho en el anterior párrafo, es inevitable –por la disposición de información estadística y por la fácil aplicabilidad de los resultados teóricos– la conexión con la literatura empírica. A lo largo del artículo, no obstante, se han intentado identificar las referencias básicas para que el lector pueda profundizar más en la literatura.

Es inevitable relacionar esta literatura con el caso español. Al respecto, en la sección 2.3, hemos aplicado el modelo teórico de tipos impositivos marginales máximos sobre la renta para comparar los óptimos con los reales. Y decimos «óptimos», pues el hecho de que el IRPF sea un impuesto descentralizado hace natural

²³ Un análisis interesante que podría realizarse y que, sin duda, contribuiría al actual debate sobre la imposición de la riqueza en España es la obtención de tipos marginales óptimos sobre el capital (rendimientos del ahorro en el IRPF más gravamen del stock en el impuesto sobre el patrimonio) por CC. AA.

el análisis por CC. AA. Obtenemos, por un lado, que hay diversidad de tipos óptimos, lo cual se debiera ver favorecido precisamente por la descentralización; por otro lado, para según qué valores de la elasticidad de respuesta, la presión fiscal que, en el margen, soportan los contribuyentes del último tramo podría llegar a estar por encima del óptimo social. Para elasticidades bajas, que pueden servir de referencia de la situación en que las posibilidades de elusión apenas existen, los niveles óptimos de imposición están bien por encima de los actuales. En la sección 3.2, ampliamos el foco para tratar también el gravamen óptimo de la renta del ahorro; éste es un tema que merece un tratamiento peculiar en nuestro país por la coexistencia de un impuesto sobre los rendimientos (en el IRPF) y de un impuesto sobre el patrimonio (Esteller-Moré, 2022).

Referencias bibliográficas

- Almunia, M., & López-Rodríguez, D. (2019). The elasticity of taxable income in Spain: 1999–2014. *SERIEs*, 10(3), 281-320.
- Atkinson, A. B. (2015). *Inequality. What Can be Done?*. Harvard University Press.
- Atkinson, A.B., & Stiglitz, J.E. (1976). The design of tax structure: direct versus indirect taxation. *Journal of Public Economics*, 6(1-2), 55–75.
- Banks, J., & Diamond, P. (2010). The Base for Direct Taxation. En J. Mirrles, S. Adam, T. Besley, R. Blundell, S. Bond, R. Chote, M. Gammie, P. Johnson, G. Myles y J. Poterba, *The Mirrlees Review: Dimensions of tax design*, Capítulo 6, (pp. 548-648). Oxford University Press.
- Blundell, R., Bozio, A., & Laroque, G. (2011). Extensive and Intensive Margins of Labour Supply: Working Hours in the US, UK and France. *American Economic Review*, 101(3), 482-486.
- Boadway, R., Marchand, M., & Pestieau, P. (1994). Towards a theory of the direct-indirect tax mix. *Journal of Public Economics*, 55(1), 71-88.
- Boadway, R., & Cuff, K. (2022). *Tax Policy. Principles and Lessons*. Cambridge Elements. Elements in Public Economics. Cambridge University Press.
- Chetty, R. (2009). Is the Taxable Income Elasticity Sufficient to Calculate Deadweight Loss? The Implications of Evasion and Avoidance. *American Economic Journal: Economic Policy*, 1(2), 31-52.
- Deaton, A. (1979). Optimally Uniform Commodity Taxes. *Economics Letters*, 2(4), 357–61.
- Diamond, P. A. (1998). Optimal Income Taxation: An Example with a U-Shaped Pattern of Optimal Marginal Tax Rates. *American Economic Review*, 88(1), 83–95.
- Diamond, P., & Saez, E. (2011). The Case for a Progressive Tax: From Basic Research to Policy Recommendations. *Journal of Economic Perspectives*, 25(4), 165-190.
- Esteller-Moré, A. (2021). Imposición óptima y descentralización fiscal: El caso del IRPF. *Investigaciones Regionales / Journal of Regional Research*, 49(1), 29-44.
- Esteller-Moré, A. (2022). Impuesto sobre el patrimonio: de la teoría a la práctica. *Economistas*, 178, 108-116.

- Feldstein, M. (1999). Tax Avoidance and the Deadweight Loss of the Income Tax. *Review of Economics and Statistics*, 81(4), 674–80.
- Gruber, J., & Saez, E. (2002). The Elasticity of Taxable Income: Evidence and Implications. *Journal of Public Economics*, 84(1), 1–32.
- Haig, R. M. (1921). *The Federal Income Tax*. Columbia University Press.
- Harju, J., & Matikka, T. (2016). The elasticity of taxable income and income-shifting: what is “real” and what is not?. *International Tax and Public Finance*, 23(4), 640–669.
- Jacobs, B. (2013). *From Optimal Tax Theory to Applied Tax Policy*. CESifo Working Paper No. 4151, Center for Economic Studies and Ifo Institute (CESifo), Munich.
- Kaldor, N. (1955). *An Expenditure Tax*. Allen and Unwin, Londres.
- Kleven, H. (2021). *The EITC and the Extensive Margin: A Reappraisal*. (NBER Working Paper No. 26405). National Bureau of Economic Research.
- Kleven, H., Landais, C., Munoz, M., & Stantcheva, S. (2020). Taxation and Migration: Evidence and Policy Implications. *Journal of Economic Perspectives*, 34(2), 119–142.
- Meade, J. (1978). *The Structure and Reform of Direct Taxation: Report of a Committee chaired by Professor J. E. Meade for the Institute for Fiscal Studies*. George Allen & Unwin, Londres.
- Mirrlees, J. A. (1971). An Exploration in the Theory of Optimal Income Taxation. *Review of Economic Studies*, 38(2), 175–208.
- Neisser, C. (2021). The Elasticity of Taxable Income: A Meta-Regression Analysis. *The Economic Journal*, 131(640), 3365–3391.
- Oates, W. E. (1999). An Essay on Economic Federalism. *Journal of Economic Literature*, 37(3), 1120–1149.
- Piketty, T., & Saez, E. (2013). Optimal Labor Income Taxation. En A. J. Auerbach, R. Chetty, M. Feldstein y E. Saez, *Handbook of Public Economics*, Vol. 5, (pp. 391–474). Elsevier.
- Piketty, T., Saez, E., & Stantcheva, S. (2014). Optimal Taxation of Top Labor Incomes: A Tale of Three Elasticities. *American Economic Journal: Economic Policy*, 6(1), 230–271.
- Pressman, M. (2018). The Ability to Pay in Tax Law: Clarifying the Concept's Egalitarian and Utilitarian Justifications and the Interactions between the Two. *N.Y.U. Journal of Legislation and Public Policy*, 21, 141–201.
- Rubio Guerrero, J. (1986). El impuesto sobre el gasto personal: información necesaria para su gestión. *Papeles de Economía Española*, 27, 303–305.
- Saez, E. (2001). Using Elasticities to Derive Optimal Income Tax Rates. *Review of Economic Studies*, 68(1), 205–09.
- Saez, E. (2002). The Desirability of Commodity Taxation Under Non-Linear Income Taxation and Heterogeneous Tastes. *Journal of Public Economics*, 83(2), 217–230.
- Saez, E., & Stantcheva, S. (2018). A simpler theory of optimal capital taxation. *Journal of Public Economics*, 162, 120–142.
- Salanié, B. (2003). *The Economics of Taxation*. MIT Press.
- Sandmo, A. (2015). The Principal Problem in Political Economy: Income Distribution in the History of Economic Thought. En A. B. Atkinson y F. Bourguignon, *Handbook of Income Distribution*, 2A, (pp. 3–65), Elsevier.
- Schanz, G. (1896). Der Einkommensbegriff und die Einkommensteuergesetze. *FinanzArchiv*, 13, 1–87.

- Simons, H. C. (1938). *Personal Income Taxation*. University of Chicago Press.
- Slemrod, J., Yitzhaki, S., Mayshar, J., & Lundholm, M. (1994). The optimal two-bracket linear income tax. *Journal of Public Economics*, 53(2), 269-290.
- Slemrod, J., & Gillitzer, C. (2014). *Tax Systems*. MIT Press.
- Slemrod, J., & Kopczuk, W. (2002). The optimal elasticity of taxable income. *Journal of Public Economics*, 84(1), 91-112.
- Sørensen, P. B. (2005). *Dual Income Taxation: Why and How?*. CESifo Working Paper No. 1551, Center for Economic Studies and Ifo Institute (CESifo), Munich.
- Tuomala, M. (1990). *Optimal Income Tax and Redistribution*. Oxford University Press.
- Tuomala, M., & Weinzierl, M. (2020). Prioritarianism and Optimal Taxation. En M. Adler y O. Norheim, *Prioritarianism in Practice*, Cap. 4, Cambridge University Press.
- Utz, S. (2002). Ability to Pay. *Whittier Law Review*, 23, 867-889.