

EVALUACIÓN DE POLÍTICAS PÚBLICAS

INTRODUCCIÓN EDITORIAL

LA justificación de la evaluación de las políticas públicas puede hacerse, al menos, desde tres dimensiones. En primer lugar, la rendición de cuentas a los ciudadanos en el uso de los fondos públicos por parte de sus gobernantes es una pieza fundamental en toda democracia. En segundo lugar, los principios de eficacia y eficiencia en el uso de los recursos públicos aparecen recogidos en nuestras normas de mayor rango (Constitución Española, Ley Orgánica de Estabilidad Presupuestaria y Sostenibilidad Financiera y Ley General Presupuestaria, por poner tres ejemplos relevantes). Por último, desde una perspectiva más pragmática, el diseño correcto de la política pública, esto es, conocer si alcanza sus objetivos y, en caso contrario, proponer su modificación o supresión exige su evaluación de manera que los programas públicos se basen en resultados y no en buenas intenciones o meras intuiciones.

El desarrollo de la evaluación de las políticas públicas, como le ha sucedido a la propia evolución de la Economía Pública en las últimas décadas como una disciplina cada vez más empírica (Martínez-Vázquez, 2018), se ha visto favorecido por dos elementos claramente relacionados. Por una parte, la disponibilidad de microdatos administrativos de carácter longitudinal. Por otra, el desarrollo de nuevas técnicas, en especial, el uso de experimentos aleatorios (*Randomized Controlled Trials* o *RCTs*), recientemente refrendadas con la concesión de varios premios nobel de Economía (Banerjee, Duflo y Kremer en 2019 y Angrist, Card e Imbens en 2021).

Si descendemos a nuestro país, desde una perspectiva institucional en términos comparados, España alcanza un nivel medio-bajo en la evaluación de políticas públicas (Stockmann, Meyer y Taube, 2020). Lo cierto es que nos falta una verdadera cultura de la evaluación como sucede en los países con una larga tradición en evaluación (Reino Unido y Estados Unidos).

Las bases de datos, en su mayoría, no están diseñadas para evaluar sino con objetivos más tradicionales de control del gasto y, en general, no se encuentran a disposición de los investigadores. En ese sentido, una interesante iniciativa es la propuesta de la Autoridad Indepen-

EL DISEÑO CORRECTO DE LA POLÍTICA PÚBLICA EXIGE SU EVALUACIÓN DE MANERA QUE LOS PROGRAMAS PÚBLICOS SE BASEN EN RESULTADOS Y NO EN BUENAS INTENCIONES O MERAS INTUICIONES

EL DESARROLLO DE LA EVALUACIÓN DE LAS POLÍTICAS PÚBLICAS SE HA VISTO FAVORECIDO POR DOS ELEMENTOS: LA DISPONIBILIDAD DE MICRODATOS ADMINISTRATIVOS DE CARÁCTER LONGITUDINAL Y EL DESARROLLO DE NUEVAS TÉCNICAS, EN ESPECIAL, EL USO DE EXPERIMENTOS ALEATORIOS

**NOS FALTA UNA
VERDADERA CULTURA
DE LA EVALUACIÓN
COMO SUCEDE EN
LOS PAÍSES CON UNA
LARGA TRADICIÓN EN
EVALUACIÓN (REINO
UNIDO Y ESTADOS
UNIDOS)**

diente de Responsabilidad Fiscal (AIReF, 2020) de llevar a cabo una estrategia nacional para la disponibilidad y el uso de los datos que contienen los registros administrativos.

En cuanto a las técnicas más novedosas de evaluación que ponen el énfasis en el establecimiento de causalidad entre los programas públicos y sus resultados e impactos, no existen especialistas en la administración pública y se encuentran al alcance solo de unos pocos académicos universitarios que las utilizan en sus investigaciones.

La situación sin embargo parece que está cambiando. Un paso relevante en la buena dirección ha sido la puesta en marcha de la División de Evaluación del Gasto Público en la AIReF en 2021, consolidando con ello la actividad que esta institución venía realizando con los llamados *Spending Reviews* desde 2017. Por otra parte, el Plan de Recuperación, Transformación y Resiliencia constituye una buena oportunidad para profundizar en la práctica evaluadora a juzgar por la forma en la que deben justificarse la percepción de los fondos asociados a ese plan.

Un número de *Papeles de Economía Española* dedicado a la «Evaluación de las Políticas Públicas» parece, en consecuencia, plenamente justificado. El que presentamos consta de doce trabajos agrupados en tres bloques; cuatro forman la parte introductoria, tres se dedican a imposición y política fiscal, y cinco dan el contenido a las políticas de gasto.

La sección introductoria comienza con un análisis comparado de la evaluación de las políticas públicas en España con ocho países caracterizados por su tradición evaluadora por parte de **José María Casado** y **Eloísa del Pino**. Los autores constatan el avance que se ha producido en nuestro país en la institucionalización de la evaluación de sus políticas públicas, así como una serie de deficiencias y retos importantes que agrupan en diferentes aspectos. En relación al marco organizativo y regulatorio de la evaluación, no existe en nuestro país una ley de evaluación de políticas, aunque sí normas que la mencionan de manera más o menos directa como sucede en la mayoría de los países. Un paso importante en esta dirección, después de la escasa exitosa experiencia de la Agencia Estatal de Evaluación de la Calidad de los Servicios y de las Políticas Públicas (AEVAL), es la puesta en marcha de la División de Evaluación del Gasto Público en la Autoridad Independiente de Responsabilidad Fiscal (AIReF) en 2021 que ha formalizado la práctica de la evaluación en esa institución independiente

**UN PASO RELEVANTE
HA SIDO LA
PUESTA EN MARCHA
DE LA DIVISIÓN DE
EVALUACIÓN DEL
GASTO PÚBLICO EN LA
AIReF EN 2021**

que llevaba desarrollándose desde 2017 y que se consolidará aún más con el papel que tendrá la institución como organismo evaluador en el Plan de Recuperación, Transformación y Resiliencia (PRTR) del Gobierno central (2021). En cuanto a la capacidad evaluadora, se pone de manifiesto la relevancia de disponer de datos administrativos en la experiencia comparada y el reto que supone para nuestro país la puesta en marcha de una estrategia nacional con ese objetivo, así como disponer de efectivos en la administración con conocimientos específicos en técnicas de evaluación. Por último, una dimensión relevante es el uso de las evaluaciones en el que queda mucho camino por recorrer si nos comparamos con países en los que la publicación de los resultados de las evaluaciones es obligatoria y se hace un seguimiento de tales resultados. En este sentido, el PRTR incluye la creación de una unidad en el Ministerio de Hacienda para el seguimiento de las propuestas de la AIReF y su incorporación al ciclo presupuestario. También ayudará a la difusión el que las evaluaciones de la AIReF sean llevadas a las conferencias sectoriales donde se reúnen los ministerios y consejerías autonómicas correspondientes.

Dado el extraordinario desarrollo producido en los últimos años en las técnicas de evaluación de impacto, era obligado dedicar un trabajo a ofrecer una panorámica de las mismas, tarea que llevan a cabo **Joaquín Artés** y **Beatriz Rodríguez-Sánchez**. A pesar de la complejidad del tema, el resultado es óptimo debido a la claridad de su contenido y al apoyo de ejemplos con el que ilustran las diferentes alternativas. Tras plantear el problema fundamental de identificación de relaciones causales en el ámbito de las políticas públicas, se aborda el método de aleatorización que permite asegurar que el grupo tratado y el grupo de control sean estadísticamente idénticos en todo aquello que no es la pertenencia al grupo tratado, lo que permite atribuir las diferencias entre los dos grupos al tratamiento o a la política aplicada. A veces la aleatorización se produce de modo natural (involuntariamente) mediante experimentos naturales. Cuando lo anterior no es posible, como sucede en la mayoría de las evaluaciones de las políticas públicas, hay que recurrir a otras técnicas estadísticas. Las más destacadas son las cuasiexperimentales que permiten identificar los efectos causales incluso cuando existen diferencias no observables entre el grupo tratado y el de control. De ellas se ocupa el trabajo en su parte final (regresión discontinua, diferencias en diferencias y variables instrumentales) destacando sus ventajas e inconvenientes en relación con las características de la política a evaluar y con la información disponible en cada caso.

EL PRTR INCLUYE LA CREACIÓN DE UNA UNIDAD EN EL MINISTERIO DE HACIENDA PARA EL SEGUIMIENTO DE LAS PROPUESTAS DE LA AIReF Y SU INCORPORACIÓN AL CICLO PRESUPUESTARIO

A VECES LA ALEATORIZACIÓN SE PRODUCE DE MODO NATURAL (INVOLUNTARIAMENTE) MEDIANTE EXPERIMENTOS NATURALES

El artículo de **Jennifer Graves** y **Eugenio Zucchelli** versa sobre un aspecto tan relevante como poco tratado en el ámbito de la evaluación de las políticas públicas, la necesidad de considerar las consecuencias inesperadas e indeseables al evaluar el efecto causal de tales políticas, que los autores ilustran mediante una serie de ejemplos para los casos de sanidad y educación. En políticas sanitarias se destaca la importancia de incluir en la evaluación los efectos sustitución entre bienes, los efectos indirectos en otros sectores y tener en cuenta las interacciones que se producen entre las legislaciones de distintos niveles de gobierno sobre una determinada materia. En políticas educativas se insiste en los incentivos que se producen en las respuestas a determinados objetivos (resultados en los exámenes) con la reducción del tiempo dedicado a desarrollar contenidos o capacidades no evaluadas tanto o más importantes. También se llama la atención sobre los problemas de implementar políticas a gran escala a partir de los resultados obtenidos al aplicar una política en un ámbito reducido, como por ejemplo la reducción del número de alumnos por aula cuyos efectos beneficiosos acaban más que contrarrestándose con una pérdida de calidad del profesorado que ha de contratarse masivamente para aplicar esa política. Como reflexión final, se aboga por una evaluación periódica de las políticas teniendo en cuenta que buena parte de esos efectos se producen pasados bastantes años desde que las políticas públicas se aplican.

Cerramos la parte introductoria con un trabajo dedicado a los desarrollos recientes en eficiencia y productividad en el ámbito del sector público a cargo de **Daniel Santín González**. A este tema ya en 1995 la revista *Papeles de Economía Española* dedicó su número 95 con el título «Sector Público y Eficiencia» también coordinado por el profesor Pedraja. Tras destacar la conexión entre eficiencia y evaluación y la importancia de mejorar aquella en nuestro sector público, dada su delicada situación financiera y la relevancia cuantitativa de los programas de gasto representativos del Estado del bienestar, el artículo expone tres desarrollos metodológicos recientes en los métodos de frontera en los que el autor ha tenido un papel protagonista. En primer lugar, las metafronteras y tecnologías de referencia, de enorme interés para estados descentralizados como sucede en nuestro país. En segundo lugar, la combinación de las herramientas de la inferencia causal con las técnicas de medición de la eficiencia basadas en el uso de fronteras. Por último, la utilización del análisis de fronteras productivas en la cuantificación de ganancias de eficiencia en procesos de fusión de unidades que prestan servicios similares. El artículo se cierra con la

propuesta de crear agencias de evaluación en el ámbito autonómico en España.

La sección de imposición y política fiscal comienza con el trabajo de **Carlos Sanz** sobre preferencias impositivas, renta de los contribuyentes y confianza en las instituciones. Se trata de un valioso análisis descriptivo, complementario de posteriores análisis causales, en el que se obtienen interesantes resultados a partir de una encuesta. Son los hogares con más renta los que expresan una mayor preferencia por subir los impuestos, lo que contradice el marco teórico tradicional en el caso de que se asocien mayores impuestos a más redistribución. El estudio muestra además que otros factores sociodemográficos y políticos correlacionan con las preferencias impositivas de los ciudadanos; en ese sentido, los hombres, las personas de más edad y los votantes de izquierda muestran una mayor preferencia por subir los impuestos. Por último, una variable clave para explicar las preferencias impositivas es la confianza en las instituciones públicas y el sistema político, contribuyendo a apoyar la literatura que identifica a ese factor de la confianza en las instituciones como un elemento crucial para el desarrollo económico y el buen funcionamiento institucional.

El artículo de **Ignacio Moral-Arce** se ocupa del impacto de las subidas de tipos reducidos en el IVA en los años 2010 y 2012. Tras repasar los estudios que analizan el impacto de los cambios en la imposición indirecta general y especial en España, utilizando tanto modelos de simulación como aproximaciones contrafactuales, analiza los efectos de esos cambios utilizando la EPF en el periodo 2007 a 2015 con técnicas de regresión en discontinuidad. Para ello distingue lo que denomina margen extensivo, número de hogares afectados, y margen intensivo o efecto sobre el gasto medio de los hogares, diferenciando el impacto de cada reforma por grupos de bienes. Un resultado interesante es el distinto efecto, por niveles de ingresos del hogar, de las reformas de 2010 y 2012 por categorías de bienes.

Al ámbito de la evaluación de la política fiscal se dedica el trabajo de **Iván Kataryniuk, Lucía López Muñoz y Javier J. Pérez** en el que, desde una perspectiva macroeconómica, se revisan los principales elementos institucionales y técnicos en la evaluación *ex ante* y *ex post* de los planes de la política fiscal. En primer lugar, se repasa el marco institucional europeo y nacional para la elaboración y supervisión de los planes fiscales, incluyendo una referencia particular al papel de los organismos de control en los mismos. En segundo lugar, se exponen

**EL PROGRAMA NGEU
ES UN EJEMPLO
PROMETEDOR
DE DISEÑO
INSTITUCIONAL
QUE INCORPORA
INCENTIVOS Y
ELEMENTOS
DE VALORACIÓN
EX ANTE Y EX POST**

los elementos técnicos básicos necesarios para una adecuada evaluación *ex ante* de los planes. En tercer lugar, se examinan los aspectos más relevantes de acuerdo con la literatura de referencia sobre la evaluación *ex post* de los planes, elemento crucial para dotar de credibilidad a los mismos y pieza esencial de la rendición de cuentas de los gobiernos frente a los ciudadanos, resaltando el caso del programa NGEU como un ejemplo prometedor de diseño institucional que incorpora incentivos y elementos de valoración *ex ante* y *ex post*.

La evaluación de las políticas públicas en la administración de justicia corre a cargo de **Juan S. Mora-Sanguinetti**, quien divide el contenido de su trabajo en cuatro secciones. En la primera, justifica la evaluación del buen funcionamiento del sistema judicial poniendo el énfasis en su relevancia económica en sus dos vertientes privada y pública. En este sentido, se insiste en la relación entre seguridad jurídica e inversión y se llama la atención sobre cómo la acción de la justicia, desde una perspectiva global, determina, en parte, el comportamiento de la productividad. En la segunda sección, tras repasar los indicadores que la literatura utiliza para evaluar la eficacia de la justicia, se decanta por su lentitud como el indicador más adecuado y repasa las fuentes de información existente para un análisis comparado entre países. En la tercera sección, a partir de las fuentes anteriores, ofrece un panorama de la situación de la justicia española y pone de relieve su elevada tasa de congestión general, así como las notables diferencias entre comunidades autónomas. Por último, mediante ejemplos de impactos económicos se analizan algunos problemas de funcionamiento de la justicia en España: marco regulatorio complejo y negativas implicaciones para la inversión de la ineficacia de la jurisdicción civil.

El artículo de **Beatriz González López-Valcárcel** constituye una excelente reflexión sobre datos y métodos de evaluación de políticas sanitarias y de salud, su organización en España y los temas relevantes para una agenda priorizada. La autora señala que, pese al enorme avance en la disponibilidad de datos, la metodología para procesarlos, para evaluar políticas públicas, está relativamente estancada. La modelización ha progresado, pero todavía falta integrar mejor los modelos epidemiológicos y económicos, así como incorporar los efectos sobre la desigualdad. En cuanto a la gestión de la pandemia, pone de manifiesto la falta de una evaluación global de la misma, que actualmente se encuentra en proceso de elaboración. La evaluación de la sanidad debe abarcar tres niveles: macro (políticas de salud y sanitarias, meso (gestión y organización de centros y servicios sanitarios) y micro (gestión

**LA MODELIZACIÓN
HA PROGRESADO,
PERO TODAVÍA FALTA
INTEGRAR MEJOR
LOS MODELOS
EPIDEMIOLÓGICOS
Y ECONÓMICOS, ASÍ
COMO INCORPORAR
LOS EFECTOS SOBRE
LA DESIGUALDAD**

clínica), siendo áreas de especial relevancia para la evaluación la digitalización de la sanidad, la gestión indirecta de la prestación sanitaria y las tecnologías de alto coste unitario, como la medicina personalizada, y finaliza reclamando, entre otras acciones, una evaluación urgente del modelo MUFACE.

El artículo de **José Manuel Cordero Ferrara, Francisco Pedraja Chaparro** y **Rosa Simancas Rodríguez** ofrece una panorámica de la evaluación en educación. Comienza justificando la importancia de la evaluación en educación y, concretamente, en los resultados educativos ya que explican una parte importante de las tasas de crecimiento de los países. A continuación, se repasan, sin un ánimo exhaustivo, algunos estudios que utilizan las técnicas de evaluación de impacto en la función de producción educativa tanto desde una perspectiva internacional como nacional. Los autores ponen de manifiesto la superioridad de los experimentos aleatorios controlados como método de evaluación cada vez más utilizados especialmente en países con mayor tradición evaluadora. Precisamente esa es la gran asignatura pendiente de España en la evaluación de políticas educativas, la utilización de experimentos aleatorios, aunque existen algunas excepciones impulsadas por diversas instituciones. Por último, a modo de ejemplo, evalúan el Programa 18-25 implementado en Extremadura consistente en la oferta de un incentivo económico a los desempleados con edades comprendidas entre 18 y 25 años para que finalizasen la enseñanza secundaria obligatoria e incrementasen con ello sus posibilidades de incorporación al mundo laboral. Mediante el uso de una regresión en discontinuidad, se demuestra que la posibilidad de obtención del título de la ESO no fue estadísticamente diferente entre el grupo de tratados y el grupo de control. En definitiva, esa política educativa, no resultó efectiva.

La evaluación de las políticas en el mercado de trabajo corre a cargo de **José María Arranz** y **Carlos García-Serrano**. Su propósito es doble, hacer una revisión de estos trabajos para el caso español y mostrar un ejemplo práctico de evaluación de alguna política realizada en el mercado de trabajo en nuestro país. En cuanto al primer objetivo, ponen de manifiesto la escasez de estudios que aplican estas técnicas apuntando un par de razones. Por una parte, la dificultad de disponer de datos adecuados (longitudinales y representativos); por otra, la simultaneidad de cambios en la normativa laboral que imposibilita la delimitación adecuada entre grupos de tratados y de control. El ejercicio de evaluación se centra en el cambio legislativo producido en 2012 por el que se aumentó la edad mínima para solicitar una

EL PROGRAMA 18-25 IMPLEMENTADO EN EXTREMADURA DEMUESTRA QUE LA POSIBILIDAD DE OBTENCIÓN DEL TÍTULO DE LA ESO NO FUE ESTADÍSTICAMENTE DIFERENTE ENTRE EL GRUPO DE TRATADOS Y EL GRUPO DE CONTROL

EL CAMBIO LEGISLATIVO PRODUCIDO EN 2012 POR EL QUE SE AUMENTÓ LA EDAD MÍNIMA PARA SOLICITAR UNA PRESTACIÓN ASISTENCIAL ILIMITADA TUVO EFECTOS SUSTANCIALES EN EL COMPORTAMIENTO DE LOS TRABAJADORES MAYORES

prestación asistencial ilimitada. Utilizando microdatos del Servicio Público de Empleo Estatal (SEPE) y mediante una técnica de diferencias en diferencias, llegan a la conclusión, análoga a la obtenida en la literatura comparada, de que la política tuvo efectos sustanciales en el comportamiento de los trabajadores mayores. Concretamente, la probabilidad de salir del desempleo hacia un empleo para las personas de 52 a 54 años que dejaron de tener acceso a los subsidios ilimitados después de que se produjo el cambio de política, fue 2,5 veces mayor que la que había antes de ese cambio.

Para finalizar, en el artículo de **Luis Ayala Cañón, Antonio Jurado Málaga y Jesús Pérez Mayo** se lleva a cabo una primera valoración del ingreso mínimo vital (IMV) puesto en marcha en junio de 2020, una importante novedad ya que se trata de la primera renta mínima de ámbito nacional en España. El tema es muy relevante dados los elevados niveles, en términos comparados europeos, de nuestro país en pobreza y desigualdad y las notables diferencias por categorías demográficas y territoriales que existen en España, donde buena parte de las competencias en esta política social se encuentra descentralizada en las CC. AA. Con la escasa información disponible y el breve periodo de desarrollo del IMV, el trabajo utiliza diversas fuentes como los microdatos de la ECV 2020, la encuesta FOESSA 2021 y los registros administrativos. Entre sus principales conclusiones se destaca que, pese a la mejora en relación con los niveles ofrecidos por los esquemas autonómicos, se observa el escaso alcance del programa para la población objetivo que se habría duplicado de haber utilizado el umbral del 40 por 100 de la mediana de la renta equivalente del hogar, el más frecuente en la literatura especializada. De los resultados también se desprende que la principal causa de denegación es el incumplimiento de los umbrales económicos de ingresos y patrimonio, por lo que resulta imprescindible una modificación del modo de determinar los recursos económicos a evaluar. Por último, se identifican problemas considerados de ineficiencia asignativa vertical al existir un significativo porcentaje de unidades que, sin pertenecer a la población objetivo, declaran recibir la prestación, lo que pone de manifiesto un evidente problema de equidad del programa.

BIBLIOGRAFÍA

- MARTÍNEZ-VÁZQUEZ, J. (2018). Perspectives on the Last Quarter Century of Research in Public Economics. *Revista de Economía Aplicada*, 76 (vol. XXVI), pp. 9-33.
- STOCKMANN, R., MEYER, W. y TAUBE, L. (eds.) (2020). *The Institutionalization of Evaluation in Europe*. Cham: Palgrave Macmillan.