



PAPELES DE ECONOMÍA ESPAÑOLA

EVALUACIÓN DE POLÍTICAS PÚBLICAS RESÚMENES EJECUTIVOS

≡ funcas

El texto íntegro del contenido de este monográfico, puede descargarlo de forma gratuita en la página web de Funcas: <https://www.funcas.es/revista/evaluacion-de-politicas-publicas/>

«Evaluación de políticas públicas»

Coordinado por Francisco Pedraja Chaparro

SUMARIO

COLABORACIONES

I. ASPECTOS GENERALES

Similitudes y diferencias en la evaluación de políticas públicas en ocho países: España en perspectiva comparada:	3	<i>José María Casado y Eloísa del Pino</i>
Métodos de evaluación de políticas públicas:	5	<i>Joaquín Artés y Beatriz Rodríguez-Sánchez</i>
Consecuencias no intencionadas de las medidas de política pública: ejemplos de sanidad y educación:	7	<i>Jennifer Graves y Eugenio Zucchelli</i>
Desarrollos recientes en eficiencia y productividad para su aplicación en el sector público:	9	<i>Daniel Santín González</i>

II. EL TRANSPORTE TERRESTRE DE PERSONAS

¿Más o menos impuestos? Renta, confianza en las instituciones y preferencias impositivas:	11	<i>Carlos Sanz</i>
Evaluación del impacto de un incremento del IVA reducido sobre el gasto de los hogares en España:	13	<i>Ignacio Moral-Arce</i>
Los planes de política fiscal: evaluación <i>ex ante</i> y <i>ex post</i> :	15	<i>Iván Kataryniuk Lucía López Muñoz y Javier J. Pérez</i>

III. GASTO PÚBLICO

La evaluación de las políticas públicas: la administración de justicia:	17	<i>Juan S. Mora-Sanguinetti</i>
---	----	---------------------------------

Evaluación de políticas. Salud y sanidad:	19	<i>Beatriz González López-Valcárcel</i>
La evaluación en el contexto educativo:	21	<i>José Manuel Cordero Ferrera Francisco Pedraja Chaparro y Rosa Simancas Rodríguez</i>
La evaluación del impacto de las políticas públicas: el caso de las políticas del mercado de trabajo:	23	<i>José M. Arranz y Carlos García-Serrano</i>
El ingreso mínimo vital: adecuación y cobertura:	25	<i>Luis Ayala Cañón Antonio Jurado Málaga y Jesús Pérez Mayo</i>

SIMILITUDES Y DIFERENCIAS EN LA EVALUACIÓN DE POLÍTICAS PÚBLICAS EN OCHO PAÍSES: ESPAÑA EN PERSPECTIVA COMPARADA

José María CASADO

Eloisa DEL PINO

AIReF

EN los últimos años, España ha avanzado significativamente en la institucionalización de la evaluación de sus políticas públicas (véase Casado y Del Pino, 2021). Aun así, el país sigue presentando retos singulares e importantes en relación con la evaluación que otros países con más tradición evaluadora ha abordado con más o menos éxito a lo largo del tiempo. Compararse con ellos es útil para calibrar los desafíos españoles en su justa medida y para aprender de sus experiencias.

El análisis de la institucionalización de la evaluación de políticas en el nivel central de gobierno en ocho países (Holanda, Suiza, Canadá, Estados Unidos, Reino Unido, Alemania, Francia y España) arroja interesantes conclusiones en relación con:

El marco organizativo y regulatorio de la evaluación

La evaluación se inscribe en un marco organizativo y regulatorio desigualmente estructurado y coordinado en los países analizados. En todos, el ejecutivo y el legislativo realizan funciones de evaluación. En Canadá y Estados Unidos existe mayor coordinación y puede hablarse de una política de evaluación en sus gobiernos federales. En el resto de los países las estructuras de evaluación del ejecutivo funcionan en gran medida de forma independiente. En el caso del legislativo, la función evaluadora está más consolidada en unos países que en otros. Los tribunales de cuentas realizan funciones de auditoría y fiscalización y, en algunos países, también llevan a cabo evaluaciones.

En España, las funciones de evaluación de políticas se ubican principalmente en el ejecutivo, aunque otras instituciones, como el Tribunal de Cuentas, realizan actividades cercanas. La puesta en marcha de la División de Evaluación del Gasto Público en la Autoridad Independiente de Responsabilidad Fiscal (AIReF) en 2021 ha formalizado la práctica de la evaluación en esa institución que llevaba desarrollándose desde 2017. El Gobierno central encarga periódicamente a la AIReF los llamados *Spending Reviews* (el segundo de los cuales constará de cinco fases para el período 2022-2026) y otras evaluaciones concretas; también doce comunidades autónomas le han encargado varias evaluaciones.

En todos los países existe algún tipo de plasmación legal de la evaluación, pero es muy variable. En algunos, la evaluación se introdujo en la constitución (Suiza y Francia), en otros existe una norma sobre el tema (Canadá y Estados Unidos); en la mayoría de los casos, la evaluación se recoge de manera más bien genérica en disposiciones generales o sectoriales.

Aunque España no cuenta con una ley de evaluación de políticas, existen normas que la mencionan de manera más o menos directa. En 2021, el Gobierno de España abrió una consulta pública para la redacción de un anteproyecto de ley. Además, el Plan de Recuperación, Transformación y Resiliencia (PRTR) del Gobierno central (2021) consolida el papel que tendrá la AIReF como organismo evaluador.

Capacidad evaluadora

En los países donde la evaluación está más consolidada, la preocupación por si las políticas se evalúan o no, ha dado paso a una creciente inquietud por la calidad y por el uso o impacto de las evaluaciones. Algunos países han avanzado en la definición de las capacidades de los evaluadores y de las unidades de evaluación; se han definido estándares de calidad para las evaluaciones; existen guías y cajas de herramientas de evaluación, pero pocos cuentan con unidades que apoyen el desarrollo de la evaluación en los ministerios. En todos los países se ha mejorado la disponibilidad y el uso de los datos administrativos.

Las evaluaciones que se realizan desde el ejecutivo se hacen internamente o contratan con terceros, en búsqueda de mayor independencia y/o calidad en porcentajes variables. En algunos países se contrata sobre todo con empresas privadas, pero en otros se recurre en mayor medida a institutos de investigación semipúblicos. La evaluación es aparentemente más frecuente en algunos sectores, como empleo o cooperación al desarrollo, o allí donde la disponibilidad de datos es mayor.

En España, la capacidad para hacer evaluaciones de los ministerios es en general limitada. Los organismos encargados de evaluar cuentan con un número de efectivos muy reducido. La AIReF busca apoyo externo (consultoras y/o universidades) para la mayoría de las evaluaciones, pero se involucra intensamente con cada uno de los equipos externos, diseñando y dirigiendo las evaluaciones y coordinando su gobernanza. El Instituto de Evaluación de Políticas ha desarrollado numerosas guías de evaluación, pero su actividad evaluadora es limitada. La AIReF (2020) propuso el diseño de una estrategia nacional que requería cambios legales para que los datos se produzcan con fines estadísticos y de investigación y evaluación de políticas, a la vez que se crean repositorios de información accesibles por las instituciones y la academia. En los últimos años se han producido algunos avances respecto a los datos, pero no son suficientes.

El uso de las evaluaciones

Existe un creciente interés por las evaluaciones en los parlamentos nacionales que se expresa de distintos modos (más menciones en discursos, solicitudes de evaluación, exigencia de que el ejecutivo respalde sus propuestas normativas con evidencia o que evalúe las políticas). En algunos países existen iniciativas sistemáticas para que todos los ministerios identifiquen los temas sobre los que requieren más evidencia y para divulgar evidencia sólida sobre políticas. Incluso en los países donde la publicación es obligatoria, las evaluaciones no siempre están disponibles; existen iniciativas para realizar un pormenorizado seguimiento de sus propuestas resultados de la evaluación.

En España, el IEPP y la AIReF publican sus evaluaciones, pero no es fácil encontrar las de los ministerios. Las evaluaciones de la AIReF han contribuido a la configuración de distintas medidas de política pública. 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 sugiere que las evaluaciones de la AIReF sean llevadas a las conferencias sectoriales, donde se reúnen los ministerios y consejerías del ramo, lo que ayudará a su difusión.

MÉTODOS DE EVALUACIÓN DE POLÍTICAS PÚBLICAS

Joaquín ARTÉS

Beatriz RODRÍGUEZ-SÁNCHEZ

Universidad Complutense de Madrid

RESULTA fundamental la evaluación de los efectos de las políticas públicas para garantizar que cumplen los objetivos fijados en su diseño o para realizar cambios en caso de no alcanzarlos y para, además, asegurar que el dinero público se usa eficientemente. Las técnicas de evaluación de impacto causal permiten medir los efectos de las políticas públicas en las variables de interés de forma que pueda evaluarse hasta qué punto la política ha logrado sus objetivos.

La evaluación de los efectos de las políticas públicas no está exenta de problemas, siendo uno de ellos, cómo diferenciar los cambios en la variable objetivo (aquello sobre lo que se quiere actuar, por ejemplo, un mejor rendimiento académico) de los cambios causados por otras variables que también puedan influir en el resultado. Si el análisis planteado no permite analizar la relación causa-efecto entre la implantación de la política pública y su resultado en la variable objetivo, o lo hace erróneamente, conducirá a conclusiones incorrectas y, por tanto, a decisiones potencialmente inapropiadas.

A pesar de los numerosos desarrollos metodológicos en la evaluación de impacto de las políticas públicas, su aplicación en el ámbito de la gestión pública en España es relativamente escaso. Por ello, y con el fin de contribuir a superar esta laguna, en este trabajo describimos los principales métodos estadísticos de evaluación de impacto causal, mostrando además su aplicación a través de distintos trabajos académicos.

Una primera parte del trabajo describe el problema analítico que representa la identificación de relaciones causales en el ámbito de las políticas públicas. El problema se plantea porque para cada individuo que incluyamos en nuestro análisis únicamente seremos capaces de observar un estado de la naturaleza: o bien ha recibido la política pública (ha sido "tratado") o no. Como no se pueden observar ambos casos de manera simultánea, el efecto causal del "tratamiento" ha de medirse recurriendo a estimaciones de lo que hubieran hecho los individuos tratados en ausencia de la política pública. En consecuencia, la principal dificultad de los métodos de evaluación de políticas públicas radica en la identificación del grupo de comparación (o control) adecuado para poder estimar de forma precisa lo que habría ocurrido con la variable objetivo en ausencia de la política pública que ha afectado al grupo de tratamiento.

Una vez descrito el problema central que supone la identificación de efectos causales, nos ocupamos del método más preciso de medición de efectos causales: la aleatorización, que consiste en elegir aleatoriamente *ex ante* qué individuos reciben el tratamiento y cuáles no. La aleatorización permite asegurar que el grupo tratado y el grupo de control son 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. Este método es habitual en el campo de las ciencias de la salud, pero menos común en el ámbito de las ciencias sociales, probablemente debido a las dificultades añadidas dadas las características de los tratamientos a analizar.

Cuando la aleatorización *ex ante* no resulta posible, se pueden identificar efectos causales recurriendo a la aleatorización que se produce involuntariamente o de forma "natural" por la forma en la que unos individuos han estado expuestos o no a una política pública. Este tipo de aleatorizaciones involuntarias reciben el nombre de experimentos naturales.

La aleatorización del tratamiento, bien sea diseñada *ex ante* o de forma natural, no es posible en la gran mayoría de casos de evaluación de políticas públicas por lo que la evaluación ha de hacerse recurriendo a

otras técnicas estadísticas, y, en particular, a métodos que permitan controlar todos aquellos factores observables que diferencien a ambos grupos. Los métodos más adecuados en estos casos son los modelos de regresión y los modelos de emparejamiento o *matching* por su denominación en inglés.

Una última parte del artículo se ocupa los métodos cuasiexperimentales, que permiten identificar efectos causales incluso cuando existen diferencias no observables entre el grupo tratado y el grupo de control. Los métodos cuasiexperimentales más conocidos son la regresión discontinua, el método de diferencias en diferencias y las variables instrumentales.

El método de regresión discontinua se utiliza para estimar efectos causales “locales” cuando la asignación al grupo tratado depende de que los individuos superen o no un determinado umbral en una variable continua (e.g. renta o nota académica), denominada variable de asignación. Los individuos que están muy cerca del umbral de corte que determina el tratamiento son presumiblemente muy parecidos, por lo que para identificar efectos causales pueden compararse individuos que se han quedado ligeramente por encima del umbral con los que están ligeramente por debajo.

El método de diferencias en diferencias resulta especialmente adecuado cuando la política pública a evaluar se aplica a individuos diferentes en diferentes momentos del tiempo. En estos casos puede compararse la diferencia en la variable objetivo entre el grupo tratado y el grupo de control antes y después de implementar una política pública. Restando las diferencias no atribuibles al tratamiento de las diferencias totales entre un grupo y otro después del tratamiento puede identificarse el efecto causal.

Por último, el método de variables instrumentales se apoya en el supuesto de que, para comprobar si los cambios en una variable causan los cambios en una variable objetivo, existe una tercera variable (el instrumento) que afecta a la probabilidad de recibir el tratamiento, pero no directamente a la variable objetivo ni a otras variables que puedan afectar también a la variable de interés. Si observamos que los cambios en la variable objetivo están asociados a cambios en el instrumento podremos identificar el efecto causal de la variable instrumentada en la variable objetivo.

En definitiva, el artículo se ocupa de describir los métodos que habitualmente se utilizan para analizar efectos causales de las políticas públicas, describiendo sus fundamentos y las ventajas e inconvenientes de cada uno de ellos en relación con la información disponible en cada caso.

CONSECUENCIAS NO INTENCIONADAS DE LAS MEDIDAS DE POLÍTICA PÚBLICA: EJEMPLOS DE SANIDAD Y EDUCACIÓN

Jennifer GRAVES

Universidad Autónoma de Madrid (UAM)

Eugenio ZUCHELLI

Madrid Institute for Advanced Study (MIAS), Universidad Autónoma de Madrid (UAM), University of Lancaster, IZA

ESTE artículo utiliza ejemplos relevantes de la literatura académica sobre política sanitaria y educativa para ilustrar la importancia de medir e interpretar correctamente el impacto causal de las políticas públicas. A través de una serie de ejemplos de actualidad, exploramos los mecanismos comunes que dan lugar a efectos inesperados y potencialmente indeseables de algunas políticas públicas. Resulta importante reconocer tales mecanismos en la estimación de los efectos casuales porque no hacerlo lleva a una comprensión errónea de los efectos completos de una política y a soluciones políticas inadecuadas.

Este estudio se divide en dos secciones principales. En primer lugar, exploramos ejemplos de la literatura académica sobre política sanitaria y consideramos las consecuencias de no tener en cuenta: i) la heterogeneidad de respuestas conductuales ante una misma política pública; ii) las posibles externalidades y efectos indirectos de las políticas públicas; y iii) la interacción entre las legislaciones de diferentes niveles de gobierno (por ejemplo, regional/local frente a nacional). En segundo lugar, utilizamos ejemplos provenientes de políticas educativas para mostrar las consecuencias de: i) no tener en cuenta la heterogeneidad de respuestas conductuales; ii) comprender incorrectamente cómo se generalizarán los resultados de los estudios científicos a contextos diferentes; y iii) no tener en cuenta los efectos de las políticas educativas en otros ámbitos como el mercado de trabajo.

Los resultados que presentamos sobre las repercusiones de las políticas sanitarias subrayan las ventajas que tiene para la evaluación de políticas públicas la combinación de diferentes fuentes de datos para tener en cuenta una amplia gama de posibles resultados y subgrupos de población heterogéneos. La calidad de los datos existentes en algunos países ha permitido analizar de forma muy precisa los efectos de políticas sanitarias relacionadas con la prohibición o reducción del consumo de determinados bienes como el alcohol o el tabaco. Estos estudios muestran que no deben considerarse los efectos sobre el uso del bien en cuestión de forma aislada, sino que también deben tenerse en cuenta en la evaluación los efectos de sustitución entre bienes similares; los efectos indirectos en diferentes sectores de la economía; y las interacciones entre la legislación local y la nacional. Además, los responsables políticos deben tener en cuenta las posibles implicaciones en materia de desigualdad. Por ejemplo, las políticas que aumentan los impuestos sobre algunos bienes (e.g. los alimentos poco saludables) podrían ser regresivas al favorecer a los consumidores de mayores ingresos que tienden a tener una dieta más saludable y a ser más sensibles a los cambios de precios.

Los ejemplos sobre políticas educativas sugieren que sería importante prever las potenciales respuestas conductuales estratégicas de los individuos afectados debido a los incentivos que introducen determinadas medidas de política educativa habitualmente utilizadas. Un ejemplo de respuesta conductual no deseada ante el establecimiento de pruebas estandarizadas para todos los alumnos es el aumento del tiempo de instrucción dedicado por los profesores a la mejora de las habilidades para hacer exámenes a costa de reducir el tiempo dedicado a impartir los contenidos. Otro ejemplo puede ser el aumento del tiempo de instrucción en aquellas áreas en las que más potencial existe para aumentar la puntuación en los exámenes estandarizados a costa de reducir el tiempo dedicado a materias potencialmente más importantes. Un segundo bloque de ejemplos de política educativa, sugieren que es importante tener en cuenta que los resultados de una

política aplicada en pequeña escala no necesariamente se mantienen al aplicarse a gran escala. El ejemplo más claro son las políticas de reducción del número de alumnos. La reducción de la ratio de alumnos por profesor es indudablemente positiva si la aplicación de la política no implica cambios en la calidad del profesorado de la contratación de nuevos profesores. Si el necesario aumento del número de profesores conlleva reducciones en la calidad del profesorado, el efecto positivo de la reducción del número de alumnos puede verse reducido o eliminado por la caída de la calidad. Finalmente, un tercer bloque de ejemplos sugiere que las políticas educativas tienen efectos en el mercado de trabajo ya que la jornada laboral de los padres y madres trabajadoras depende de la jornada escolar. Los cambios que afectan a la jornada escolar tienen, en consecuencia, efectos en el mercado de trabajo que, a menudo, son ignorados.

Concluimos instando a los investigadores y a los responsables políticos a que presten atención a la forma en que los diversos mecanismos esbozados en este documento podrían aplicarse en el diseño y elaboración de nuevas políticas públicas. También sugerimos que sería igualmente importante comprender y reconocer que muchos de los efectos secundarios de las políticas que discutimos aquí no fueron descubiertos por los investigadores hasta varios años después de que la política entrara en vigor. Esto pone de manifiesto, una vez más, la importancia de la evaluación periódica de las políticas y la necesidad de que los investigadores y los responsables políticos colaboren en las primeras fases de su aplicación.

DESARROLLOS RECIENTES EN EFICIENCIA Y PRODUCTIVIDAD PARA SU APLICACIÓN EN EL SECTOR PÚBLICO

Daniel SANTÍN GONZÁLEZ

Universidad Complutense de Madrid

Desde un punto de vista legal, tanto la Constitución Española como la ley presupuestaria recogen el mandato de eficiencia en la gestión de los recursos públicos. Por ello, conviene volver a recordar la importancia de llevar a cabo una evaluación sistemática de la eficiencia y la productividad con la que actúan los servicios públicos mediante el uso de fronteras productivas. Año tras año la oferta de servicios públicos que se realiza mediante unidades independientes de producción pública, como colegios, hospitales, juzgados de primera instancia, etc. consume un importante porcentaje de los recursos de las administraciones públicas con el objetivo de entregar servicios de calidad a los ciudadanos.

Las principales ventajas de llevar a cabo la evaluación del desempeño de estos servicios públicos mediante fronteras productivas y competencia por comparación son numerosas. En primer lugar, ayuda a definir el cumplimiento de objetivos y a clarificar las responsabilidades de las organizaciones. En segundo lugar, permite una mejor coordinación de las políticas dentro de cada administración. También mejora la transparencia, ofreciendo indicadores de desempeño de los servicios públicos a lo largo del tiempo. Finalmente, realizar *benchmarking* permite identificar a las mejores prácticas, aprender de su gestión y promover la introducción de mejoras e innovación en la entrega de los servicios.

En España, el primer Plan de Actuación del *Spending Review* desarrollado por AIREF marcó un punto de inflexión al incluir, por primera vez de manera explícita en un texto de una institución pública, la posibilidad de realizar *benchmarking* mediante el uso de fronteras productivas, lo que AIREF denomina “el cálculo de fronteras eficientes de producción, que determinan un *benchmarking*”. En numerosas ocasiones las evaluaciones de eficiencia y productividad no se llevan a cabo por parte de las administraciones públicas o por parte de otros agentes argumentando que no existe información acerca de todas las variables con las que sería necesario contar. En este sentido es importante señalar que constantemente existirá información e indicadores no disponibles y siempre será posible perfeccionar las medidas ya existentes. Este hecho genera un círculo vicioso ya que los indicadores no se generan porque no existe evaluación y la evaluación nunca empieza porque se considera que no existen los indicadores adecuados. El resumen es que para romper este círculo es imprescindible la voluntad política de comenzar a evaluar. Queda pendiente en España desarrollar evaluaciones sistemáticas que implementen este objetivo de *benchmarking* para lo cual se proponen tres ideas.

En primer objetivo es la evaluación global de la industria. En el caso del sector público español entendemos por industria al conjunto de unidades productivas que producen un servicio en un determinado sector (centros de salud, institutos de secundaria, etc.) que está bajo la jurisdicción de una misma administración que normalmente será la comunidad autónoma. La evaluación se debe centrar entonces en analizar la eficiencia de cada industria y ver cómo evoluciona su productividad a lo largo del tiempo. A continuación, el objetivo sería comparar el desempeño de la industria con otras industrias semejantes. Para ello se puede utilizar el concepto de meta-frontera que consiste en trazar una única frontera a partir de las mejores prácticas de las unidades productivas pertenecientes a las distintas industrias. Alternativamente se puede utilizar una tecnología de referencia para comparar industrias.

Con esta idea se trata de recuperar la idea de “Federalismo de Laboratorio”, propuesta por el profesor Oates, como una importante ventaja del sistema autonómico español todavía no bien explotada. Un sistema descentralizado como el español permite probar y experimentar distintas alternativas para entregar un servicio público.

En segundo lugar, una vez lograda la institucionalización de la medición de los cambios en productividad de los distintos servicios públicos, el siguiente objetivo sería el análisis causal acerca de qué cambios en la gestión permiten mejorar el desempeño. Para ello, una línea actual de investigación consiste en combinar las herramientas tradicionales de inferencia causal con las técnicas de medición de la eficiencia basadas en el uso de fronteras.

En tercer lugar, las técnicas de fronteras productivas pueden ser utilizadas para analizar las ganancias esperadas en eficiencia; efecto aprendizaje, efecto armonía y efecto de escala, como consecuencia de fusionar unidades que gestionan el mismo servicio. Las aplicaciones potenciales de esta metodología son numerosas tanto adentro de grandes organizaciones, por ejemplo, departamentos adentro de una universidad, o servicios como por ejemplo el de la fusión de municipios.

Se concluye que para llevar a cabo una adecuada evaluación de la eficiencia y la productividad de las unidades productivas encargadas de gestionar los servicios públicos sería imprescindible crear Agencias Autónomas Independientes de Evaluación de todos los Programas y Servicios Públicos que dependan de uno o varios parlamentos autonómicos. En España, las CC. AA. tienen las competencias en sanidad y educación, los dos principales motores de gasto público "en especie". Por tanto, la administración autonómica, que es quien asigna el presupuesto, debería ser también la primera interesada y encargada en evaluar "sus propias industrias" contando con agencias independientes.

¿MÁS O MENOS IMPUESTOS? RENTA, CONFIANZA EN LAS INSTITUCIONES Y PREFERENCIAS IMPOSITIVAS

Carlos SANZ

Banco de España

¿ CUÁLES son las preferencias impositivas de los ciudadanos? En este artículo, apporto nueva evidencia sobre las preferencias de la población española sobre el tamaño del sector público, basándome en los datos obtenidos en una encuesta *online* llevada a cabo con Mónica Martínez-Bravo (CEMFI) en colaboración con la empresa de datos YouGov.

En primer lugar, documento que los individuos que viven en hogares de más renta expresan una mayor preferencia por subir los impuestos y son más proclives a expresar una opinión positiva sobre ellos. En particular, los resultados indican que la preferencia por subir impuestos aumenta de forma aproximadamente lineal por renta. Los coeficientes son cuantitativamente importantes: por ejemplo, los individuos en el quintil más rico tienen una preferencia sobre impuestos casi un punto (sobre diez) más alta que los individuos en el quintil más pobre. En segundo lugar, estudio qué otros factores sociodemográficos y políticos correlacionan con las preferencias impositivas de los ciudadanos. Encuentro que los hombres y las personas de más edad y más nivel educativo, así como los votantes de izquierdas (PSOE y UP) expresan una mayor preferencia por subir los impuestos. En concreto, las mujeres tienen una preferencia significativamente menor por subir impuestos que los hombres: 0.26 puntos (en una escala 0-10). Respecto a los jóvenes (<31 años), los individuos de edad media (de 31 a 65 años) no muestran una diferencia significativa, mientras que los mayores (>65 años) tienen una mayor preferencia por aumentar los impuestos. La diferencia es cuantitativamente relevante: 0.57 puntos (en la escala 0-10). También hay diferencias significativas por nivel educativo: los resultados indican que los individuos con menos educación (menos de secundaria) tienen una menor preferencia por subir los impuestos. Respecto a los votantes del PSOE, los votantes de UP tienen una preferencia 0.63 puntos (sobre 10) mayor por subir los impuestos, mientras que los de Cs, PP y VOX tienen una preferencia 1.24, 2.17 y 2.79 puntos menor, respectivamente. Las diferencias por nivel de ingresos en las preferencias impositivas, no obstante, se mantienen aun después de tener en cuenta todos estos factores. Por último, encuentro que una variable clave para explicar las preferencias impositivas es la confianza en las instituciones públicas y el sistema político: las personas con más confianza expresan una mayor preferencia por subir los impuestos y son más proclives a mencionar opiniones positivas sobre los mismos. Esta correlación es cuantitativamente importante, estadísticamente significativa y está presente después de controlar por el resto de factores de renta, sociodemográficos y políticos. El coeficiente estimado implica que un aumento de un punto (sobre 10) en la confianza en el sistema político correlaciona con un aumento de 0.29 puntos (sobre 10) en la preferencia por subir los impuestos.

Este artículo contribuye a la literatura que estudia los determinantes del tamaño del sector público. El clásico modelo del votante mediano predice que una mayor desigualdad debería llevar a más impuestos y mayor redistribución (Meltzer y Richard, 1981). Intuitivamente, dado que el votante mediano es más pobre (antes de impuestos) que la media, esto dará lugar a más impuestos y redistribución. Aumentos de la desigualdad (normalmente asociados a una mayor diferencia entre la media y la mediana) conducirán de este modo a un mayor nivel de impuestos. Un supuesto clave de estos modelos es que los individuos votan de acuerdo a su propio interés y que este depende exclusivamente de la renta. De este modo, el modelo supone que los individuos más ricos votarán en favor de menores impuestos que los pobres, contrariamente a la evidencia que el presente artículo presenta para España.

La literatura posterior a Meltzer y Richard (1981) ha intentado analizar la relación predicha por el modelo de que mayor desigualdad antes de impuestos debería llevar a mayor redistribución. Relacionado con el presente artículo, hay evidencia de que la confianza de los individuos en las herramientas redistributivas

de la política económica influye en las preferencias sobre redistribución. En este sentido, Kuziemko *et al.* (2015), encuentran que dar información detallada sobre la desigualdad de renta afecta significativamente a la preocupación por la desigualdad, pero apenas afecta a las preferencias sobre impuestos y transferencias (con la notable excepción del impuesto de sucesiones o *estate tax*). Encuentran que esto se debe a la falta de confianza en el gobierno y en la desconexión percibida entre los problemas sociales y las políticas públicas dirigidas a abordarlas.

El presente artículo contribuye también a la literatura que estudia la relevancia de la confianza en las instituciones. Existe considerable evidencia de que la confianza es crucial para el desarrollo económico y el buen funcionamiento institucional (ver, por ejemplo, Algan y Cahuc, 2010, y Acemoglu *et al.*, 2020; ver también la revisión de esta literatura de Sorribas-Navarro, 2021, en esta misma revista). Los resultados del presente artículo apuntan también a una fuerte correlación de la confianza en el sistema político y las instituciones con las preferencias con el tamaño del sector público.

Referencias

ACEMOGLU, D., CHEEMA, A., KHWAJA, A. I. y ROBINSON, J. A. (2020). Trust in state and nonstate actors: Evidence from dispute resolution in Pakistan. *Journal of Political Economy*, 128(8), pp. 3090–3147.

ALGAN, Y. y CAHUC, P. (2010). Inherited trust and growth. *American Economic Review*, 100(5), pp. 2060-2092.

KUZIEMKO, I., NORTON, M. I., SAEZ, E. y STANTCHEVA, S. (2015). How elastic are preferences for redistribution? Evidence from randomized survey experiments. *American Economic Review*, 105(4), pp. 1478-1508.

MELTZER, A. y RICHARD, S. (1981). A rational theory of the size of government. *Journal of Political Economy*, 89, pp. 914–927.

SORRIBAS-NAVARRO, P. (2021). Corrupción y confianza en los políticos en España. *Papeles de Economía Española*, (168), pp. 157-191.

EVALUACIÓN DEL IMPACTO DE UN INCREMENTO DEL IVA REDUCIDO SOBRE EL GASTO DE LOS HOGARES EN ESPAÑA

Ignacio MORAL-ARCE

Instituto de Estudios Fiscales

ESTE trabajo analiza el impacto que un incremento del tipo de gravamen en el IVA reducido en los años 2010 y 2012 de 1 y 2 puntos, respectivamente, ha tenido sobre el gasto de los hogares en España, empleando una aproximación de evaluación con contrafactuales, y usando la metodología de regresión en discontinuidad para la estimación.

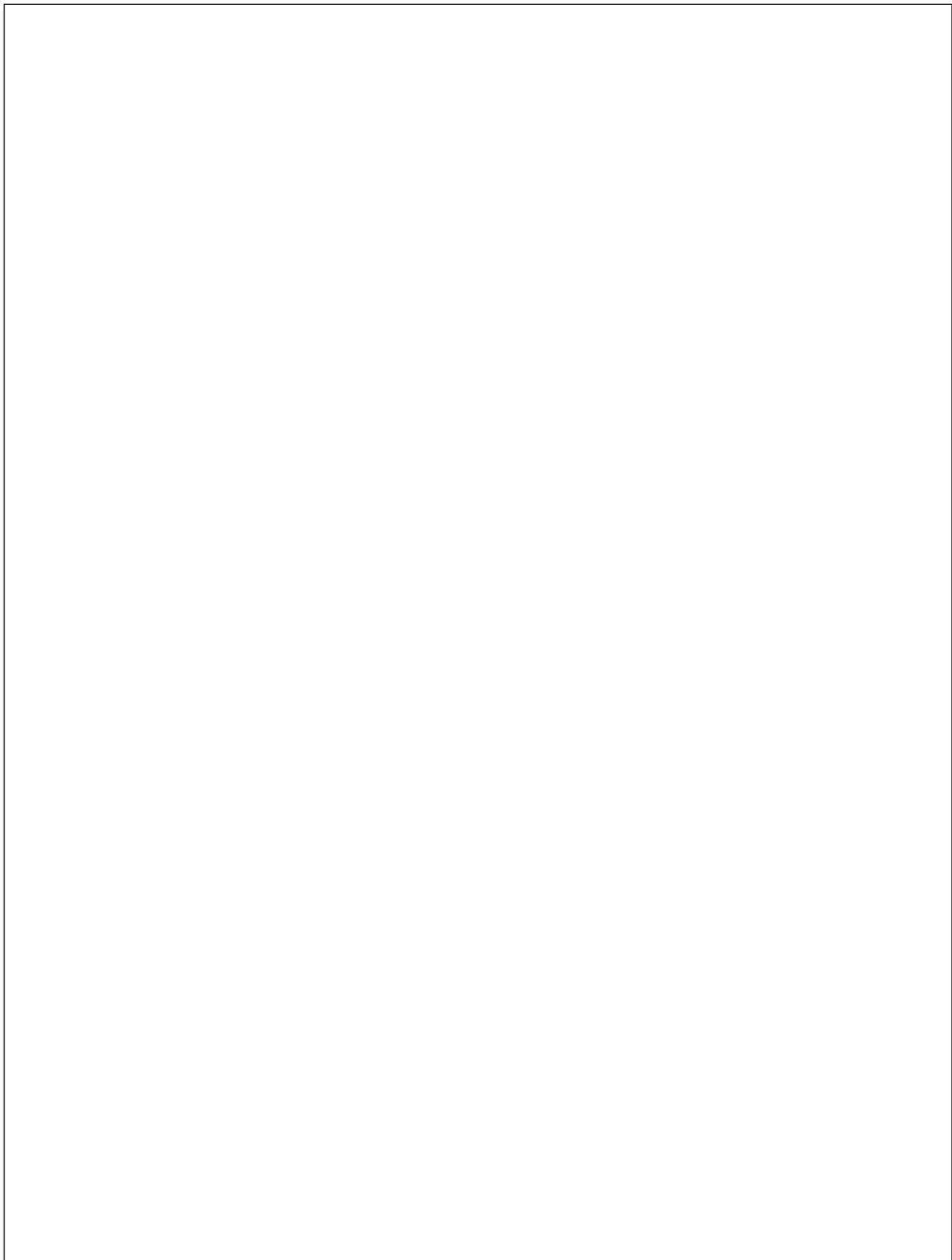
Inicialmente se ofrece un análisis de las principales modificaciones normativas que ha sufrido el IVA reducido en el periodo 2001 hasta 2020, tanto de entradas y salidas de bienes y servicios como cambios de gravamen.

A continuación, se muestran las principales investigaciones que hay en España sobre el impacto de cambios en impuestos indirectos, IVA e impuestos especiales, con investigaciones que tienen un enfoque más macro, utilizando modelos de simulación, hasta aproximaciones más microeconómicas, empleando técnicas de contrafactuales.

Con datos mensualizados de la *EPF* para el período 2007–2015, se evalúa tanto el cambio en el número de hogares que consumen bienes o servicios afectados por el IVA reducido –margen extensivo– como el efecto sobre el gasto medio de los hogares –margen intensivo–, empleando el método de regresión en discontinuidad. Además, las estimaciones se desagregan para diferentes tipos de bienes y servicios y de hogares.

Para el margen extensivo, los resultados muestran que la subida de dos puntos de 2012 reduce ligeramente el número de hogares que gastan en alimentos y bebidas, mientras que el porcentaje de hogares que realizan gastos en hostelería aumenta en 2010. En el margen intensivo, la reforma de 2012 afecta negativamente al gasto medio de los hogares en este tipo de bienes y servicios, influyendo especialmente en una reducción en el gasto en bebidas no alcohólicas, restauración y hostelería.

Por lo que respecta al análisis de heterogeneidad dependiendo del nivel de ingresos del hogar, el resultado más relevante es el efecto negativo de la subida de gravamen de 2010 en los hogares de renta baja, con una reducción del número de consumidores en alimentos y bebidas, mientras que la reforma de 2012 afectó negativamente a rentas altas en estos tipos de bienes, mientras que estudiando el gasto medio de los hogares, la subida de 2010 afecta principalmente a rentas altas en el gasto en bebidas no alcohólicas y transporte, mientras que la de 2012 influye, sobre todo, en hogares con rentas bajas con una reducción del gasto medio en diferentes bienes y servicios.



LOS PLANES DE POLÍTICA FISCAL: EVALUACIÓN EX ANTE Y EX POST

Iván KATARYNIUK

Lucía LÓPEZ MUÑOZ

Javier J. PÉREZ

Banco de España

La rendición de cuentas de los gobernantes frente a los ciudadanos en el uso de los fondos públicos es una de las piezas fundamentales de la democracia, que en el caso de la política fiscal se refleja en el proceso de elaboración y ejecución de los presupuestos públicos dentro de los parámetros establecidos por la ley, y en un contexto de transparencia y control externo. La rendición de cuentas requiere una clara comprensión del impacto económico de la acción discrecional de los gobiernos, manifestada en las políticas de ingresos y gastos públicos. Esto ha llevado gradualmente en las últimas décadas, y de manera más intensa en Europa desde mitad de la década de los años noventa del pasado siglo, al establecimiento de procedimientos e instituciones presupuestarias más sólidas, a la publicación de datos sobre la ejecución presupuestaria cada más frecuentes y sujetos a mejores estándares técnicos, y al desarrollo de herramientas cada vez más sofisticada para realizar el seguimiento y la previsión de la actividad de las administraciones públicas.

Las medidas de política fiscal se concretan, en condiciones normales, en el ciclo de elaboración y aprobación del presupuesto anual que fija la legislación vigente, que puede incluir la definición de algunas actuaciones con compromisos de gasto en un horizonte plurianual (por ejemplo, en las áreas de infraestructuras o defensa). En situaciones extraordinarias, además, los gobiernos pueden aprobar paquetes fiscales fuera de las fechas habituales, típicamente en respuesta a situaciones de crisis, como en marzo de 2020, en el contexto de la respuesta a la emergencia por la COVID-19, en diciembre de 2011, como reacción a las tensiones en los mercados de deuda soberana europea, o en noviembre de 2008, como respuesta a la crisis financiera global. Algunas actuaciones presentan, por su propia naturaleza, un carácter multianual, y afectan a varios ejercicios presupuestarios. Un ejemplo reciente de esto último, son los proyectos para desarrollar en varios años mediante los fondos europeos aprobados bajo el paraguas del *NGEU*.

El impacto y la eficacia de los planes presupuestarios, a nivel agregado, se encuentra íntimamente ligado a la transparencia en su definición, lo que viene condicionado, además, por la comunicación eficaz de los mismos y la credibilidad que los agentes económicos les asignen, a la fiabilidad de su implementación siguiendo las líneas aprobadas, así como a la rendición de cuentas, esto es, a la justificación de las desviaciones observadas *ex post* con respecto a los planes definidos *ex ante*, si estas se produjeran.

En las fases de anuncio y aprobación, los agentes tienen que evaluar los planes, esto es, examinarlos y asignar probabilidades de impacto a los mismos, bajo distintos supuestos sobre su definición y su ejecución, y tomar decisiones de consumo, inversión u oferta y demanda de trabajo, condicionadas a las sendas hipotéticas resultantes. En esta fase, en el caso español, cuentan con los informes correspondientes de la AIReF, así como con la valoración que suelen proporcionar o publicar, distintos analistas y organismos independientes.

En la fase de implementación, los analistas tienen que evaluar, en tiempo real y con información incompleta y típicamente de publicación muy retrasada, el ajuste de las iniciativas anunciadas a actuaciones concretas, esto es, validar si lo que los gobiernos anunciaron se corresponde con lo que los gobiernos están haciendo, y actualizar las decisiones tomadas o adoptar otras nuevas.

Finalmente, *a posteriori*, una vez conocida la información necesaria, es posible evaluar el ajuste de los planes anunciados y aprobados a la implementación *ex post*, y en caso de que se hubieran producido

desviaciones, analizar las razones de estas. Por ejemplo, si estas se debieron a errores en las previsiones macroeconómicas y financieras en las que se basaron las estimaciones de ingresos públicos, a errores en la estimación de los multiplicadores fiscales asociados a las medidas, o a cambios discrecionales no anticipados en los planes de gasto y su composición (por ejemplo, ligados al ciclo electoral), entre otras posibles causas. En un proceso iterativo, el análisis de las desviaciones presupuestarias y sus motivos resulta crucial, más allá de su interés académico, para la credibilidad que los agentes económicos asignen a los nuevos anuncios de planes fiscales, lo que determinará su reacción ante estos.

En este artículo se revisa la experiencia disponible sobre estos temas. 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 los elementos técnicos básicos necesarios, en nuestra opinión, para una adecuada evaluación *ex ante* de los planes. En tercer lugar, se exponen los aspectos más relevantes que resalta la literatura de referencia sobre la evaluación *ex post* de los planes, elemento crucial para la forja de la credibilidad de estos 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: LA ADMINISTRACIÓN DE JUSTICIA

Juan S. MORA-SANGUINETTI¹

Banco de España e ICAM

ESTE artículo argumenta por qué es importante evaluar el funcionamiento del sistema judicial, expone cuáles son las mejores medidas de eficiencia judicial y las metodologías disponibles, cuantifica la situación de la justicia española y, finalmente, analiza los impactos económicos de que la justicia no funcione adecuadamente, proponiendo aplicaciones prácticas.

La primera parte (sección I) del artículo justifica la importancia de la evaluación y el análisis de la justicia y su funcionamiento. Si bien la importancia jurídica del buen funcionamiento del sistema judicial es bien conocida, parece necesario incidir en su importante relevancia económica. Más en detalle, desde un punto de vista jurídico, el sistema judicial es un instrumento fundamental de garantía de derechos y de libertades. Esta garantía se manifiesta en dos vertientes diferentes: en primer lugar, porque los ciudadanos y las empresas se enfrentan, en el ámbito privado, a otros sujetos (privados) que resultan ser oportunistas e incumplen sus obligaciones. El sistema judicial actuaría en ese ámbito como disciplinante de esas relaciones. En segundo lugar, esos mismos ciudadanos o empresas pueden verse enfrentados a la actuación incorrecta por parte de los poderes públicos. El sistema judicial es, en ese caso, una barrera o límite contra la violación de los derechos de propiedad por parte de la acción gubernamental.

Por el lado de la importancia del buen funcionamiento de la justicia para la economía, cabe recordar que, ciertamente, esas dos perspectivas (la privada y la pública) tienen una alta relevancia desde el punto de vista del análisis económico y ya fueron puestas en valor por Coase (1960) y Acemoglu y Johnson (2005). La calidad de la justicia se relaciona, en fin, con una mayor seguridad jurídica, lo que sostiene la inversión (Dejuán y Mora-Sanguinetti, 2021), una más fácil combinación de factores productivos y la toma de decisiones empresariales. Vista de forma global, la acción de la justicia determina parcialmente el comportamiento de la productividad. Además, de forma específica, más allá de su impacto en las decisiones de inversión, la ineficacia de la justicia afecta, por ejemplo, a los mercados de crédito o al mercado de vivienda (reduciendo el peso del mercado de alquiler).

La segunda parte (sección II) estudia cómo se puede proporcionar una medición adecuada de la eficacia (o ineficacia) del sistema judicial. Destaca que la bibliografía a nivel internacional reduce la evaluación normalmente a tres parámetros solamente: la lentitud de su funcionamiento, su coste o su predictibilidad. Esas tres medidas, sin embargo, ceden ante la primera: la lentitud de la justicia parece ser el indicador global más adecuado de calidad o eficacia. En la misma sección, el artículo repasa las fuentes de información estadística que pueden ser útiles para realizar la evaluación: a nivel internacional (cubriendo también el caso español) pueden destacarse los informes y bases de datos de la OCDE, la CEPEJ y el proyecto Doing Business del Banco Mundial (esta última discontinuada). A nivel nacional, España de forma individual posee muy buenas estadísticas judiciales, haciendo públicas algunas series de datos desde 1995.

La tercera parte (sección III) argumenta que, con esas bases, es posible calcular numerosos resultados estadísticos que ponen de manifiesto la situación de la justicia en España en términos objetivos. Antes del inicio de la pandemia por la COVID-19, en 2019, la justicia española, en todo lo relacionado con conflictos entre empresas y ciudadanos, tenía capacidad para resolver solamente 1 de cada 1,8 casos pendientes de su pila de trabajo. Durante la pandemia, en 2020, esa relación empeoró hasta una proporción de 1 de cada 2. Junto a ese hallazgo general, España muestra una alta variabilidad local en la tasa de congestión

¹ Las opiniones y las conclusiones recogidas en este artículo representan las ideas del autor, con las que no necesariamente tiene que coincidir el Banco de España o el Eurosistema.

judicial. La congestión es mayor en las provincias de Andalucía, Cataluña y Madrid. Por tanto, cabe señalar que el problema se agrava en las provincias mediterráneas y en Madrid, que concentran la mayor parte de la población española.

Por último, la sección IV desarrolla algunas implicaciones prácticas y ejemplos (en cuanto a los impactos económicos) del funcionamiento de la justicia en España. Primero, se analiza uno de los motivos por los que la justicia puede tornarse ineficaz y este es el de encontrarse un marco regulatorio “complejo”. Los resultados apuntan a que una complejidad «relacional» más alta, en otras palabras que las normas realicen referencias a muchas otras normas para lograr completar su contenido, dificulta o enlentece la acción judicial. Segundo, se proporcionan detalles de la relación entre inversión y justicia, poniendo de manifiesto que un incremento de la ineficacia judicial en la jurisdicción civil tiene implicaciones negativas para la inversión a nivel de empresa en España.

EVALUACIÓN DE POLÍTICAS. SALUD Y SANIDAD

Beatriz GONZÁLEZ LÓPEZ-VALCÁRCEL

Universidad de Las Palmas de Gran Canaria

La evaluación de políticas de salud es particularmente compleja porque la salud se genera fundamentalmente fuera del sector sanitario. Durante la pandemia, las políticas sanitaria y económica han sido más interdependientes que nunca. En este artículo reflexionamos sobre datos y métodos de evaluación de políticas sanitarias y de salud y sobre su organización en España, y apuntamos temas relevantes para una agenda priorizada.

En sanidad, evaluaciones *ex post* de las intervenciones y políticas públicas son útiles, pero cada vez resulta más necesaria la evaluación prospectiva *ex ante* para facilitar la toma de decisiones fundamentada en datos.

Ha habido un enorme avance del conocimiento durante los años de pandemia 2020-2021, en parte gracias al uso de datos masivos compartidos, de acceso abierto y en tiempo real, a la disponibilidad de registros sanitarios digitales poblacionales, de datos longitudinales y de *big data* (vs. datos estructurados). La inteligencia artificial ha de utilizar esos datos de forma complementaria a la inteligencia humana.

Frente al enorme avance en la disponibilidad de datos, la metodología para procesarlos para evaluar políticas públicas está relativamente estancada. Desde marzo 2020 se han elaborado modelos COVID que desde los modelos epidemiológicos *compartimentales* más simples, han ido añadiendo complejidad de grano fino, incorporando información micro sobre movilidad, interacciones sociales y políticas públicas (sintetizadas con índices estandarizados como el *Stringency Index* de Oxford). Los modelos epi-económicos son de enorme utilidad para tomar decisiones. Sirven para explicar el pasado, predecir a corto plazo, retratar el presente (*nowcasting*) y avanzar hipótesis plausibles sobre el futuro (con enorme incertidumbre). Igual que la gestión de la pandemia, los modelos deben ser evaluados retrospectivamente por su capacidad predictiva. La modelización ha progresado, pero todavía falta integrar mejor los modelos epidemiológicos y económicos, así como incorporar los efectos sobre la desigualdad.

A grandes rasgos, la gestión de la pandemia abarca: 1) las medidas para limitar los contagios; 2) la disposición y organización de recursos y servicios sanitarios para prevenir y tratar la enfermedad; 3) las medidas de carácter económico para minimizar los efectos de la pandemia sobre el tejido productivo y sobre las economías familiares.

Hoy por hoy no hay estudios rigurosos que permitan establecer la dominancia de la estrategia de supresión sobre la de contención de la COVID, o viceversa.

Los métodos de investigación cualitativa son muy útiles en la evaluación de la respuesta a la pandemia. Las denominadas *after-action reviews* (AAR) o "revisiones pos actuación" pueden servir para identificar las lecciones aprendidas y diseñar constructivamente respuestas más eficaces sin necesidad de identificar culpables. En España se reclamó una evaluación global de la gestión de la pandemia desde las primeras oleadas y el Consejo Interterritorial del Sistema Nacional de Salud (SNS) hizo el encargo en septiembre de 2021. El informe se encuentra en proceso de elaboración.

La institucionalización de la evaluación de tecnologías sanitarias en España ha avanzado en los últimos años, pero todavía no está completada. La arquitectura institucional es compleja y poco resolutive y no existen criterios explícitos de priorización de grupos de pacientes o enfermedades, aunque se reconoce la necesidad de priorizar y se acepta el criterio de coste-efectividad. La Red Española de Agencias de Evaluación de Tecnologías Sanitarias y Prestaciones del Sistema Nacional de Salud (RedETS) se constituyó formalmente en

2012. Siete comunidades autónomas (Andalucía, Aragón, Canarias, Cataluña, Galicia, Madrid, País Vasco) y el Instituto de Salud Carlos III colaboran en red y elaboran informes no preceptivos sobre determinadas tecnologías, destacando las no farmacológicas. La RedETS aporta valor al SNS y aprovecha economías de alcance y de red. En 2021, un grupo de académicos y sanitarios reclamaba una agencia independiente de evaluación de prácticas y políticas sanitarias, al estilo de la NICE británica, con capacidad legal para recomendar e inducir a los gobiernos a cumplir las recomendaciones o, alternativamente, a explicar por qué no las cumplen. La iniciativa llegó a la Comisión de Sanidad del Congreso de los Diputados en noviembre de 2021 con una proposición no de ley, pero no pasó de ahí.

Durante la pandemia quedó en suspenso el coste-efectividad como criterio de decisión, tanto por imperativo moral como por la apabullante evidencia de que prácticamente todo gasto sanitario en prevención y tratamientos COVID es dominante, teniendo en cuenta los costes (y ahorros de costes) indirectos por la vía del impacto en la economía. Tanto el rastreo de contactos (que forma parte de una estrategia más general de Test, Rastreo y Aislamiento (TRA)) como las vacunas COVID parecen ser altamente coste-efectivos, aunque la evidencia es volátil en la medida en que la pandemia cambia, y mejora nuestro conocimiento sobre ella.

La evaluación en sanidad debe abarcar los tres niveles: macro (políticas de salud y sanitarias), meso (gestión y organización de centros y servicios sanitarios) y micro (gestión clínica). España, por su descentralización sanitaria, acumula una enorme cantidad de experiencias organizativas y de gestión cuya evaluación permitiría la mejora por comparación y el escalado en todo el territorio.

La agenda de evaluación de intervenciones sanitarias de la AIREF constituye un buen punto de partida para evaluaciones macro.

Áreas de especial relevancia para la evaluación son la digitalización de la sanidad, la gestión indirecta de la prestación sanitaria y las tecnologías de alto coste unitario, como la medicina personalizada. Hay que evaluar las intervenciones económicas y fiscales sobre la salud poblacional, y analizar la actividad sanitaria con perspectiva de sistema basado en el valor. El disponer de criterios sociales de disposición a pagar por un Año de Vida Ajustado por Calidad, con posibilidad de diferenciación por grupos de pacientes o patologías, es una cuestión abierta.

El modelo Muface requiere evaluación urgente. En la agenda inmediata deberían entrar también las políticas para hacer frente a las consecuencias a largo plazo de la COVID sobre la salud (también mental) de la población y la organización de los recursos necesarios para abordar el problema, así como su financiación. Las desigualdades en el uso de servicios sanitarios, que pueden derivar en desigualdades en salud, han de ser evaluadas sin demora.

LA EVALUACIÓN EN EL CONTEXTO EDUCATIVO

José Manuel CORDERO FERRERA

Francisco PEDRAJA CHAPARRO

Rosa SIMANCAS RODRÍGUEZ

Universidad de Extremadura

EL trabajo ofrece una panorámica de la evaluación en educación tanto en el ámbito internacional como nacional. Comienza justificando la importancia de la evaluación en educación y, concretamente, en los resultados educativos dada su importancia en la explicación en las tasas de crecimiento de los países. A continuación, se repasan distintos ejemplos, sin un ánimo exhaustivo, de algunos estudios que utilizan las diferentes técnicas de evaluación de impacto al analizar el efecto de los factores que habitualmente forman parte de la función de producción educativa.

Existe un amplio consenso entre los investigadores en considerar los experimentos aleatorios controlados como el método ideal para evaluar una política educativa. Ejemplos clásicos son el programa STAR que analiza el efecto de la reducción del número de alumnos por aula o el Proyecto Perry sobre la efectividad de los programas de educación preescolar, realizados ambos en EE. UU. Su uso se ha ido extendiendo incluso en algunos países menos desarrollados para evaluar, por citar algunos casos, la entrega de dinero o comidas gratuitas en las escuelas para incentivar la asistencia a clase o la incorporación de un profesor de refuerzo.

Precisamente la utilización de experimentos aleatorios controlados es la gran asignatura pendiente de España en la evaluación de políticas educativas. En el trabajo se exponen algunas pocas excepciones, como el estudio piloto desarrollado por el Banco de España para evaluar programas de educación financiera, el proyecto Erasmus+ de la Unión Europea sobre evaluación en educación cívica, o las iniciativas de la Fundación la Caixa evaluando programas de innovación educativa y poniendo en contacto centros escolares con equipos de investigadores universitarios especializados en este tipo de técnicas.

A veces la aleatorización se produce de modo natural por el modo en que se lleva a cabo una política. Existen en nuestro país algunos estudios que aprovechan esta circunstancia. Así sucedió con la ampliación de la gratuidad de la educación infantil a los tres años con la LOGSE aplicada en las CC. AA. en distintos momentos o la influencia diversos factores en la repetición de curso, entre otros, el mes de nacimiento.

Habitualmente no es posible realizar estos experimentos aleatorios por razones muy diversas de tipo ético, económico o, simplemente, porque las políticas educativas no se diseñan con la intención de ser evaluadas. En esos casos, los investigadores acudimos a las denominadas técnicas cuasiexperimentales, como son las de diferencias en diferencias, variables instrumentales o regresión en discontinuidad. Todas pretenden imitar las condiciones de un experimento aleatorio a partir de los datos observados después de la política educativa con el fin de evitar los posibles sesgos derivados de la existencia de endogeneidad en los datos.

El método de diferencias en diferencias requiere disponer de información sobre los grupos de tratados y de control antes y después de la aplicación de la política educativa por lo que exige la utilización de datos longitudinales que suelen estar disponibles en países desarrollados en los que periódicamente se realizan evaluaciones nacionales a gran escala o registros administrativos de carácter censal. La literatura ofrece trabajos en los que aplican esta técnica a datos de sección cruzada cuando se dispone de, al menos, dos observaciones diferentes de la variable de interés para cada individuo de los grupos de tratamiento y control. Ello hace posible su utilización con datos procedentes de bases internacionales como PISA o TIMSS aunque los autores cuestionan la validez de los resultados obtenidos con esta estrategia.

La regresión en discontinuidad también es muy exigente en la información requerida en su aplicación, especialmente, alrededor del punto de corte en la variable que determina la separación entre el grupo tratado y el de control. Esta es la razón por la que la mayor parte de los trabajos utilizan datos administrativos. En cambio, el método de variables instrumentales, del que se ofrecen también algunos ejemplos de estudios en el campo de la educación, no requiere tanta información pero en cambio es fundamental la identificación del instrumento que pueda ser considerado como una alteración exógena del proceso educativo como puede suceder cuando se produce un cambio en la asignación de recursos escolares entre distritos o zonas derivados de modificaciones en la legislación educativa.

En nuestro país, los trabajos que han utilizado estas metodologías cuasiexperimentales son todavía escasos, lo que en buena medida se explica por las deficiente información existente y su falta de disponibilidad generalizada. En todo caso, con la participación de España en las principales evaluaciones internacionales y la realización de evaluaciones censales periódicas en las Comunidades Autónomas, estas limitaciones se han reducido, lo que ha permitido que se lleven a cabo trabajos que han explotado los microdatos que estas bases proporcionan para evaluar, por ejemplo, los efectos de políticas como la repetición de curso, la ampliación gratuita de la educación infantil a los tres años, los programas de bilingüismo o el uso creciente de las nuevas tecnologías en las aulas.

Por último, el artículo presenta un ejemplo de evaluación de una política educativa relacionada con el mercado de trabajo, el Programa 18-25 implementado por la Comunidad autónoma de Extremadura cuyo objetivo era elevar el nivel educativo de los desempleados extremeños entre 18 y 25 años ofreciendo el incentivo económico de 1.000 euros a aquellos obtuvieran el título de la enseñanza secundaria obligatoria (ESO) con el fin de incrementar sus posibilidades de incorporación al mercado de trabajo. 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 DEL IMPACTO DE LAS POLÍTICAS PÚBLICAS: EL CASO DE LAS POLÍTICAS DEL MERCADO DE TRABAJO

José M. ARRANZ

Carlos GARCÍA-SERRANO

Universidad de Alcalá

UN aspecto relevante de las políticas públicas dirigidas al mercado de trabajo es saber si son efectivas o no. Por un lado, las políticas pasivas del mercado de trabajo (que básicamente comprenden el sistema de prestaciones por desempleo, tanto el nivel contributivo como el asistencial) tratan de compensar la pérdida de ingresos sufrida por las personas trabajadoras como consecuencia de la pérdida del empleo, que puede afectar negativamente a la búsqueda y la disponibilidad de las personas perceptoras. Por otro lado, las políticas activas del mercado de trabajo (la orientación y la intermediación laboral, los programas de formación y los incentivos a la contratación y el emprendimiento, principalmente) están diseñadas para compensar estos posibles efectos adversos a la vez que tratar de conseguir lo más rápidamente posible la vuelta al empleo de las personas paradas y reducir los problemas de acceso y permanencia en el empleo a los que se enfrentan ciertos grupos desfavorecidos. Para saber si estas políticas son efectivas (si realmente producen algún efecto y ayudan a las personas trabajadoras que participan en ellas) es necesario evaluarlas.

El objetivo de este artículo no es revisar las técnicas cuantitativas de evaluación de impacto, sino más bien pasar revista a cómo se han aplicado en el pasado para llevar a cabo una correcta evaluación del impacto de algunas políticas públicas vinculadas al mercado de trabajo. En concreto, el propósito del artículo es doble. Por un lado, mostrar cómo se han utilizado los métodos de evaluación de impacto en los análisis del mercado de trabajo con algunos trabajos que emplean alguno de los métodos para estimar el impacto de determinados programas en los resultados del mercado de trabajo español. Por otro lado, presentar una aplicación que analiza el impacto de un cambio legislativo en el SPD y sus efectos sobre las salidas desde el desempleo de los trabajadores mayores. Este cambio legislativo se produjo en España en julio de 2012, cuando se elevó de 52 a 55 la edad para tener derecho a una prestación asistencial (o subsidio por desempleo) ilimitada. Para ello se utilizan los microdatos de los perceptores de prestaciones por desempleo puestos a disposición por el Servicio Público de Empleo Estatal (SEPE).

En cuanto a la revisión de la literatura que se ha realizado, esta muestra que la evaluación de las políticas del mercado de trabajo todavía no está muy extendida en el caso de España. El número de trabajos que llevan a cabo evaluaciones de impacto es bastante limitado, de modo que para algunos programas de las políticas activas o pasivas solo hay unos pocos trabajos que utilicen las técnicas de evaluación. Una de las razones puede ser la dificultad de disponer de datos adecuados (sobre todo, longitudinales), que además sean representativas, para poder llevar a cabo los análisis pertinentes. Otra es que en muchas ocasiones la simultaneidad de cambios en la normativa laboral hace casi imposible delimitar tanto los grupos de tratamiento como los grupos de control adecuados.

Por lo que respecta a la reforma que se analiza en el artículo, el aumento de 52 a 55 años de la edad mínima para solicitar el 'subsidio para mayores de 52/55 años' en 2012 redujo el período de derecho ilimitado para el grupo de personas de 52 a 54 años en el momento de la pérdida del empleo, proporcionando un entorno ideal para una evaluación cuasi-experimental de la influencia de la reducción de la duración potencial de las PA en las salidas del desempleo hacia el empleo. Nuestro trabajo empírico sigue la línea de la literatura que evalúa el impacto de la extensión / reducción de la duración potencial de las prestaciones por desempleo y sus efectos sobre las transiciones desde una PA hacia el empleo utilizando un enfoque de diferencias en diferencias (o dobles diferencias). Los resultados obtenidos sugieren que los cambios de este importante parámetro del periodo de derecho de las prestaciones provocan efectos sustanciales en el

comportamiento de los trabajadores mayores. En particular, 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 antes del cambio, una vez que se controla la situación económica experimentada durante el período posterior a la reforma y las diferencias de las tasas de salida medias entre los grupos de edad de 52 a 54 y de 49 a 51 años. Este hallazgo se mantiene cuando utilizamos un grupo de control alternativo (personas de 55-57 años). Estos resultados son similares a los hallazgos de la literatura empírica que apuntan a la reducción del empleo y el aumento del desempleo y la inactividad de los trabajadores mayores afectados por programas de extensión de las prestaciones por desempleo.

EL INGRESO MÍNIMO VITAL: ADECUACIÓN Y COBERTURA

Luis AYALA CAÑÓN

UNED, Equalitas

Antonio JURADO MÁLAGA

Jesús PÉREZ MAYO

Universidad de Extremadura

EL sistema de garantía de ingresos mínimos en España está formado por un mosaico de prestaciones muy variadas, con un entramado variado de prestaciones dirigidas a colectivos y necesidades específicas. Su objetivo es garantizar unos ingresos mínimos, suficientes para reducir los niveles de pobreza y desigualdad. La preocupación de la Unión Europea por estos dos problemas es creciente y es particularmente importante en España por su elevada magnitud en el contexto comparado.

En España, son dos los niveles administrativos implicados en este sistema: la Administración Central y las administraciones autonómicas. Las rentas mínimas de inserción (RMI) fueron implantadas por las comunidades autónomas entre finales de los años ochenta y el ecuador de los noventa. La Unión Europea ya invitaba a principios de aquella década a los países miembros a desarrollar programas de rentas mínimas que garantizaran el derecho a unos recursos básicos que asegurasen una vida digna a todas las personas. Estos se fueron desarrollando en varios países que hasta entonces no disponían de un programa nacional, con cierta convergencia en la definición de esquemas que trataban de unir dotación económica e inserción social.

Estos nuevos programas de renta mínima han sido una de las principales innovaciones de política social desde las últimas décadas del siglo XX, contribuyendo a la reestructuración de la mayoría de los Estados de bienestar europeos. Su extensión ha sido considerablemente más limitada en España, con un diagnóstico conocido sobre sus problemas: su complejidad, al desarrollarse como respuesta a necesidades muy específicas, la persistencia pese a ello de lagunas y obstáculos en el acceso, altos niveles de inequidad por categorías demográficas y territorios, una baja adecuación y una limitada eficacia en la activación de las personas beneficiarias.

Esto hace que la respuesta al aumento de las necesidades sociales en las fases recesivas sea menos eficaz que la de otros sistemas de garantía de ingresos europeos. Los niveles de adecuación de las prestaciones durante la Gran Recesión fueron claramente inferiores a los de otros países europeos. En 2020, España se enfrentó a una nueva crisis sin haber completado el diseño de la última red de protección social. El drástico aumento de las necesidades de los hogares por el impacto económico de las restricciones de la actividad, las citadas lagunas del sistema, con colectivos desprotegidos, y la complejidad de los procedimientos administrativos para conseguir las prestaciones lo hacían necesario.

El dilema era cómo combinar una respuesta inmediata al aumento de la vulnerabilidad económica de los hogares con la creación de una nueva prestación de carácter permanente, que contribuyera a cerrar la última red de protección económica. En junio de 2020 se puso en marcha el ingreso mínimo vital (IMV), con una novedad importante respecto al conjunto de prestaciones que articulaban previamente el sistema de garantía de ingresos, al tratarse de la primera renta mínima de ámbito nacional.

Desde su inicio, su eficacia ha sido cuestionada por el reducido número relativo de unidades beneficiarias. La premura de su puesta en marcha, sin contar con una red de gestión de las solicitudes diseñada específicamente para hogares con características distintas de los perceptores de prestaciones de la Seguridad Social y sin un sistema de información que permitiera monitorizar adecuadamente la evolución del pro-

grama han hecho que el cumplimiento de algunos de esos objetivos, sobre todo la cobertura de los hogares vulnerables, haya sido hasta ahora muy limitado.

Este trabajo supone una primera valoración de algunos de sus resultados en términos de eficacia y eficiencia asignativa. Un aspecto, sin duda, positivo, es la constatación de una mejora en los indicadores de adecuación respecto al nivel previamente ofrecido por los esquemas autonómicos. Los indicadores, sin embargo, todavía quedan lejos de los de programas similares en los países europeos de mayor renta, especialmente en el caso de los hogares de mayor dimensión.

Con la escasa información disponible, dado el breve periodo de desarrollo del IMV, el trabajo ha intentado también medir el grado de falta de respuesta (*non take-up*) de la prestación. Mediante los microdatos de la ECV 2020 y de la encuesta FOESSA 2021, junto a registros administrativos, se han obtenido algunos resultados relevantes. En primer lugar, el criterio de pobreza usado por la Administración para determinar la población objetivo en el momento de aprobar el IMV se asemeja al 30 por 100 de la mediana de la renta equivalente del hogar. Si se hubiera utilizado el umbral del 40 por 100 de la mediana, más frecuente en la literatura especializada, la población objetivo casi se habría duplicado.

Atendiendo a la falta de acceso, destaca notablemente el limitado alcance de la medida en sus primeros 18 meses. Solamente ha llegado a un tercio de los beneficiarios previstos y casi a la mitad de los hogares esperados, todo ello en el mejor escenario de identificación de la población objetivo. Si la meta principal es acabar con la pobreza severa, se puede afirmar que, por el momento, es una medida con una eficacia limitada. Una de las razones principales es la deficiencia en la información sobre el IMV.

De los resultados se desprende también que la principal causa de denegación es el incumplimiento de los umbrales económicos de ingresos y patrimonio. Parece necesaria la revisión de la manera de determinar los recursos económicos que se deben evaluar. Por último, se constata también la presencia de problemas de ineficiencia asignativa vertical, dado el porcentaje de unidades no pertenecientes a la población objetivo que declaran recibir la prestación. Este resultado revela la existencia no solo de amplios márgenes de mejora en la asignación de estas prestaciones, sino también de un problema de equidad.

El IMV, en síntesis, es una buena política, necesaria para cubrir algunas de las lagunas históricas del sistema de garantía de ingresos y reducir su fragmentación territorial, pero está afectada todavía por problemas importantes de diseño y gestión que limitan su eficacia. Aunque algunas de estas deficiencias son propias de cualquier instrumento de reciente creación, hay otras que deben ser revisadas y reformuladas para conseguir una mayor reducción de los altos niveles de pobreza de la sociedad española.

