
Evaluación y calidad de la gestión pública: reformas basadas en las mejores prácticas

Jorge Onrubia-Fernández

Instituto Complutense de Estudios Internacionales (ICEI-UCM) y FEDEA

2023 / II

EUSKO JAURLARITZA



GOBIERNO VASCO

**EKONOMIA ETA OGASUN
SAILA**
Ekonomia eta Plangintza Zuzendaritza

**DEPARTAMENTO DE ECONOMÍA
Y HACIENDA**
Dirección de Economía y Planificación

IKERKETAK EKONOMIAZ

Ikerketak-Ekonomiaz Eusko Jaurlaritzako Ekonomia eta Plangintza Zuzendaritzaren argitalpen elektronikoa da. EAEko ekonomian arreta jartzen duten zuzendaritza barruko zein kanpoko ekonomisten lanak biltzen ditu, bai ekonomia analisiaren ikuspegitik bai politika publikoenetik. Argitalpenaren helburua gure ekonomiaren alderdi garrantzitsuenen ezagutzan eta diagnostikoan sakontzea da, eta eztabaida zein politika ekonomikoaren arrazoibide onena eskaintzea. Bertan adierazten diren iritziak egileenak bakarrik dira eta ez dute zertan Eusko Jaurlaritzako Ekonomia eta Ogasun Sailaren jarrera ofizialarekin bat etorri.

Ikerketak-Ekonomiaz es una publicación electrónica de la Dirección de Economía y Planificación del Gobierno Vasco que recoge trabajos, tanto de economistas de la propia dirección como del exterior, que versan sobre la economía vasca desde la perspectiva del análisis económico y de las políticas públicas. Su objetivo es profundizar en el conocimiento y diagnóstico de aspectos relevantes de nuestra economía y propiciar el debate y la mejor fundamentación de la política económica. Los puntos de vista expresados en ella son de exclusiva responsabilidad de las personas autoras, que no tienen por qué coincidir con la posición oficial del Departamento de Economía y Hacienda del Gobierno Vasco.

DICIEMBRE 2023



Dirección de Economía
y Planificación

Resumen

El documento ofrece un panorama sobre el estado actual de la evaluación de políticas públicas en España, con el propósito de mostrar las implicaciones que, desde un punto de vista institucional, tiene su implantación operativa. Para ello, se abordan varios temas, inspirados, en buena medida, en los trabajos publicados en el volumen 103 de *Ekonomiaz*, "Calidad de las finanzas públicas: El papel de la evaluación". En primer lugar, se trata el papel de la evaluación y su relación con la mejora de la calidad de las finanzas públicas, con especial atención a los condicionantes y requisitos de su institucionalización, desde una perspectiva integral de la evaluación, conectada con la gestión pública. En segundo lugar, se presentan los principales instrumentos de los que hoy dispone la evaluación de programas y proyectos públicos, con sus potencialidades, limitaciones y exigencias en materia de datos. En tercer lugar, se revisan, de forma sintética, una serie de experiencias internacionales, consideradas como mejores prácticas por la literatura sobre evaluación pública. En cuarto lugar, tras repasar los antecedentes, el documento aborda la situación actual de la evaluación pública en España, para, tras una valoración crítica de la recientemente aprobada Ley de Institucionalización de la Evaluación de Políticas Públicas, esbozar los principales elementos que, a nuestro juicio, deberían contribuir positivamente a alcanzar una implantación efectiva de la evaluación pública en España.

Palabras clave: *evaluación de políticas públicas, calidad de las finanzas públicas, institucionalización de la evaluación, gestión pública, eficiencia del gasto público.*

Códigos JEL: *D04, H43, H50, H72, H83*

Laburpena

Dokumentuak Espainiako politika publikoen ebaluazioaren egungo egoerari buruzko panorama eskaintzen du, ikuspegi instituzional batetik bere ezarpen operatiboak dituen inplikazioak erakusteko. Horretarako, hainbat gai jorratzen dira, neurri handi batean, Ekonomiaz 103. zenbakian, "Finantza publikoen kalitatea: Ebaluazioaren zeregina", argitaratutako lanetan oinarrituta. Lehenik eta behin, ebaluazioaren zeregina eta finantza publikoen kalitatearen hobekuntzarekin duen lotura jorratzen dira, eta arreta berezia jartzen da hura instituzionalizatzeko baldintzatzaile eta eskakizunetan, kudeaketa publikoarekin lotutako ebaluazioaren ikuspegi integral batetik. Bigarrenik, programa eta proiektu publikoen ebaluazioak gaur egun dituen tresna nagusiak aurkezten dira, bere ahalmen, muga eta datuen arloko eskakizunekin. Hirugarrenik, ebaluazio publikoari buruzko literaturak jardunbide hoberentzat jotzen dituen nazioarteko esperientzia batzuk sintetikoki berrikusten dira. Laugarrenik, aurrekariak errepasatu ondoren, dokumentuak Espainiako ebaluazio publikoaren egungo egoera jorratzen du eta onartu berri den Politika Publikoen Ebaluazioaren Instituzionalizazio Legearen balorazio kritikoa egin ondoren, gure ustez Espainian ebaluazio publikoa eraginkortasunez ezartzen lagundu beharko luketen elementu nagusiak zirriborratzen ditu.

Hitz gakoak: *politika publikoen ebaluazioa, finantza publikoen kalitatea, ebaluazioaren instituzionalizazioa, kudeaketa publikoa, gastu publikoaren eraginkortasuna.*

JEL sailkapena: *D04, H43, H50, H72, H83*

Abstract

This research paper provides an overview of the current state of public policy evaluation in Spain, aiming to highlight the technical implications that its operational implementation has from an institutional perspective. To achieve this, various topics are addressed, largely inspired by the articles included in Volume 103 of *Ekonomiaz*, "Quality of Public Finances: The Role of Evaluation". Firstly, the role of evaluation and its relationship with improving public finances quality is discussed, paying special attention to the conditions and requirements for its institutionalization from a comprehensive perspective connected to public management. Secondly, the main tools available for the evaluation of public programs and projects are presented, along with their potential, limitations, and data requirements. Thirdly, a series of international experiences, considered best practices by the literature on the evaluation of public policies, are synthetically reviewed. Fourthly, after reviewing the background, the document addresses the current state of evaluation in Spain. Following a critical assessment of the recently approved Spanish Law on Institutionalization of Public Policy Evaluation, the paper outlines the main elements that, in our view, should positively contribute to achieving effective implementation of public evaluation in Spain.

Keywords: *Public policy evaluation, public finances quality, institutionalization, public management, public spending efficiency.*

JEL Codes: *D04, H43, H50, H72, H83*

Índice

Resumen ejecutivo.....	7
1. Introducción.....	12
2. La evaluación como instrumento de supervisión y fomento de la calidad de las finanzas públicas	18
2.1. El papel de la evaluación de políticas públicas	19
2.2. Ámbito funcional de la evaluación pública	23
3. Herramientas al servicio de la evaluación económica de las políticas públicas	27
3.1. El Análisis Coste-Beneficio (ACB)	28
3.2. El Análisis Coste-Efectividad (ACE).....	31
3.3. La medición de la eficiencia en las unidades gestoras	34
3.4. La evaluación de impacto	38
4. El marco institucional de la evaluación de políticas públicas	44
4.1. Aspectos teóricos del diseño institucional.....	45
4.2. Mejores prácticas internacionales en evaluación pública.....	48
4.2.1. Alemania	48
4.2.2. Australia	49
4.2.3. Canadá	51
4.2.4. Estados Unidos	54
4.2.5. Francia	57
4.2.6. Países Bajos	58
4.2.7. Reino Unido	59
4.2.8. Suecia	60
5. La institucionalización de la evaluación de políticas públicas en España: Valoración de la situación actual	64
5.1. Los antecedentes de la evaluación de políticas públicas en España.....	65
5.1.1. La evaluación de políticas públicas en la Administración General del Estado	66
5.1.2. La evaluación de políticas públicas de las comunidades autónomas	69
5.2. La evaluación de políticas en el Plan de Recuperación, Transformación y Resiliencia, y la Ley 27/2022, de institucionalización de la evaluación de políticas públicas en la Administración General del Estado	70
6. Líneas para una reforma eficaz de la evaluación pública en España.....	75
Referencias	83

Resumen ejecutivo

Este documento tiene por objetivo abordar diferentes facetas del tema de la calidad de las finanzas públicas, poniendo el foco principal en la evaluación de programas y políticas públicas. El propósito es cubrir los diferentes aspectos que, en nuestra opinión, constituyen los elementos claves de la relación entre evaluación y calidad de las finanzas públicas. En concreto, se abordan las cuestiones relacionadas con la institucionalización de la evaluación, la utilización de las evaluaciones como instrumentos de selección de los proyectos y programas que conforman las políticas públicas, bajo criterios de eficiencia económica y eficacia, así como las potencialidades de la evaluación como instrumento para mejorar la gestión pública.

En buena medida, las valoraciones y propuestas presentadas en este documento recogen, sintéticamente, las contempladas en los artículos publicados en el volumen 103 de *Ekonomiaz-Revista Vasca de Economía*, "Calidad de las finanzas públicas: El papel de la evaluación", I/2023.¹

Los procesos de consolidación fiscal suelen aparejar medidas de ajuste adoptadas con urgencia, consecuencia de los rápidos deterioros experimentados por los ingresos públicos y, consecuentemente, por el déficit público. Esto hace que los ajustes fiscales consistan, casi siempre, en recortes lineales de los créditos presupuestarios inicialmente aprobados, o sobre los montantes sin ejecutar, cuando no la supresión de recursos destinados a programas de carácter plurianual. Además, la dificultad de modificar normativas legales que reconocen derechos individuales a percibir prestaciones con cuantías preestablecidas hace que una parte importante de los recortes en los programas se concentre en el gasto público correspondiente a la provisión de servicios públicos, como la educación, la sanidad o los servicios sociales, ya que resulta más sencillo modificar o restringir las condiciones de acceso a los mismos, cuando no directamente suprimirlos. Además, la menor contestación social hace que la inversión pública sea la que soporte, con habitualidad, el mayor impacto en los ajustes, a pesar de su contrastado efecto sobre el crecimiento económico. Esta forma de afrontar los procesos de consolidación, sin tener en cuenta el grado de eficiencia concreto con el que se ejecuta cada programa de gasto, ni su efectividad medida en términos de impacto económico y social, acarrea costes de bienestar nada despreciables.

¹ <https://www.euskadi.eus/web01-a2reveko/es/k86aEkonomiazWar/ekonomiaz/getPubl?idPubl=100>

Sin embargo, una parte importante de esos costes de bienestar podrían evitarse o, al menos, mitigarse, si las reducciones del gasto fuesen selectivas y se adoptasen a partir de la información proporcionada por la evaluación económica de los programas y proyectos públicos. Los ahorros que podría generar una evaluación rigurosa del gasto público convierten a esta en una herramienta con enorme valor para afrontar estrategias de consolidación fiscal, contribuyendo además a obtener nuevas ganancias de bienestar social en el medio y largo plazo. En este sentido, el aprovechamiento de la evaluación en el diseño de estrategias de reestructuración de las políticas públicas proporcionaría un “doble dividendo”: primero, porque permitiría reducir el gasto menos productivo en términos sociales, facilitando además la supresión de aquellas actividades y programas asentados en el presupuesto por razones fundamentalmente de cálculo político; y segundo, porque una fracción de los recursos presupuestarios liberados por los ajustes adoptados podría reorientarse hacia proyectos y programas socialmente más rentables.

En el caso de España, con una trayectoria poco satisfactoria en cuanto a disciplina fiscal, pensamos que abordar un plan de reforma racionalizadora del gasto público, en términos de eficiencia y efectividad, debería ser una meta de la agenda política estatal, autonómica y local, imprescindible. De hecho, un plan de este tipo resulta esencial para poder valorar la sostenibilidad a medio y largo plazo de las finanzas públicas y en definitiva del Estado del Bienestar, muy afectado por las tensiones de gasto asociadas con el envejecimiento poblacional. Un plan como este debería contar con varias etapas: la primera debería consistir en una revisión sistemática de todos los programas presupuestarios en vigor, incluyendo análisis rigurosos de la estructura y composición de gasto destinado a cada uno, valorando las alternativas de implementación disponibles; la segunda debería incluir un análisis del grado de desempeño alcanzado por estos programas en sus años de funcionamiento; la tercera, aportar una valoración de los programas en términos de impacto económico y social, a partir de una evaluación económica basada en evidencias empíricas sostenidas por indicadores contrastables de cumplimiento de sus objetivos. Posteriormente, la revisión debería dar lugar a un informe en el que se mostrase cómo ha cambiado la estructura de programas de gasto a lo largo del tiempo, qué niveles de gobierno han puesto en marcha cambios de gestión y han practicado ajustes presupuestarios de manera motivada, valorando en todos ellos sus efectos en relación con los objetivos perseguidos.

En relación con el alcance de la evaluación de políticas, programas y proyectos públicos, se defiende en el trabajo la necesidad de delimitar previamente su ámbito

funcional, destacando la importancia que deben tener el enfoque económico y el análisis cuantitativo. Aunque la transversalidad es un rasgo distintivo presente en las actuaciones públicas, explicado por su amplia diversidad sectorial, evaluar los resultados alcanzados por las mismas, la eficiencia con la que se utilizan los recursos o el impacto cuantitativo sobre el bienestar social requieren una medición cuantitativa sujeta a estándares rigurosos y aceptados de forma unánime por los especialistas y la literatura. A lo anterior hay que añadir la importancia que tiene para la evaluación la disponibilidad de buenos datos, pues resultan esenciales para la medición rigurosa de costes y beneficios, así como para la aplicación de las técnicas cuantitativas adecuadas.

La diferenciación entre la evaluación ex ante y ex post es esencial para diseñar un buen sistema de evaluación institucional, pues ambas contribuyen, secuencialmente, a mejorar los procesos de asignación de recursos públicos y a la gestión pública en su conjunto. La evaluación ex ante es consustancial a la elección eficiente de las alternativas de diseño de las políticas, seleccionando los programas y proyectos que ofrecen una mayor rentabilidad social. En el caso de la evaluación ex post, además de contribuir a detectar los problemas tanto del diseño de los programas y proyectos como de su gestión, en un marco de transparencia, los resultados proporcionados constituyen un elemento indispensable para un adecuado funcionamiento de la rendición de cuentas y la exigencia de responsabilidades sobre la gestión realizada.

En relación con las técnicas de evaluación disponibles, el documento ofrece un inventario actualizado de las más habituales, señalando su idoneidad para las diferentes categorías de la evaluación, así como apuntando sus potencialidades y limitaciones. En el trabajo se diferencian tres categorías principales de la evaluación: a) los métodos de selección de proyectos de gasto público, integrantes fundamentales de la evaluación ex ante; b) las técnicas de medición de la eficiencia, orientadas tanto a la gestión de los programas y proyectos de gasto público como al funcionamiento de las unidades organizativas que los implementan; y c) las técnicas utilizadas para evaluar el impacto de esos programas y proyectos en el bienestar de los ciudadanos, así como el de las normativas reguladoras puestas en marcha por los gobiernos. En el primer bloque aparecen las dos técnicas tradicionalmente empleadas: el Análisis Coste-Beneficio (ACB) y su simplificación en forma de Análisis Coste-Efectividad (ACE). El segundo de los bloques incluye los métodos de medición de la eficiencia productiva, empleados como técnicas dirigidas a medir el comportamiento de eficiencia de las unidades encargadas de la provisión de bienes y servicios públicos. El tercer bloque se dedica a los métodos utilizados para realizar

evaluaciones económicas del impacto de programas, servicios públicos o regulaciones legales.

Por obvio que pueda parecer, la simple aplicación de las técnicas no garantiza que los resultados de las evaluaciones prácticas vayan a rendir los efectos de mejora de la calidad de las finanzas públicas que se espera. En este sentido, asegurar este aprovechamiento es el papel fundamental que debe desempeñar el marco institucional de la evaluación. La importancia de la institucionalización de la evaluación se aborda, en detalle, en la cuarta sección del trabajo. En primer lugar, se ofrece una revisión de los postulados teóricos que se han ido proponiendo en la literatura sobre el tema, con especial atención al diseño de la arquitectura organizativo-institucional del sistema de evaluación. Se pone el acento en la independencia entre los organismos responsables de la evaluación y los centros gestores evaluados, así como la transparencia con la que deben difundirse los resultados de las evaluaciones realizadas. La determinación de los destinatarios de los informes de evaluación, el diseño de los procedimientos para su difusión y la fijación de las implicaciones derivadas de los resultados son objeto también de tratamiento. Finalmente, basado en los trabajos previos de Albi y Onrubia (2016, 2018), se recoge, de forma sintética, un panorama de las mejores prácticas de modelos institucionales de evaluación de las políticas públicas, incluyendo las experiencias actualizadas de países como Alemania, Australia, Canadá, Estados Unidos, Francia, Países Bajos, Reino Unido y Suecia.

En la quinta sección, el estudio presta atención al caso de España. Para ello, se parte de los antecedentes de la evaluación de políticas públicas, desde la aprobación de la Constitución Española de 1978 hasta su situación actual en el Estado y en las comunidades autónomas. En esta sección, se dedica un apartado a la Ley 27/2022, de 20 de diciembre, de institucionalización de la evaluación de políticas públicas en la Administración General del Estado, recientemente aprobada, en desarrollo del Componente 11 del Plan de Recuperación, Transformación y Resiliencia, dedicado a la Modernización de las Administraciones Públicas. Además de exponer sintéticamente sus principales contenidos, el trabajo refleja una valoración crítica de la misma, apoyada en los análisis contenidos en los trabajos de De la Fuente y Novales (2023), publicado en el número 103 de *Ekonomiaz*, así como en el ensayo reciente de Novales (2023). En resumen, la valoración realizada de la Ley 27/2022 concluye que el marco legal de institucionalización de la evaluación aprobado en España (a falta de la creación de la Agencia de Evaluación prevista en la ley) es una norma llena de ambigüedades, al no aclarar quién debe ejercer las responsabilidades

técnicas ni institucionales de la evaluación, ni cuál es la estructura sistemática en la que se aplicarán los diferentes tipos de evaluación, de acuerdo con su funcionalidad específica, dentro del modelo integral con el que se quiere contar. Además, la ley tampoco concreta las metodologías estándar de evaluación de políticas públicas que se deben utilizar en cada caso, ni quiénes deben ser los destinatarios de los informes de evaluación, ni las consecuencias en términos de rendición de cuentas y de exigencia de responsabilidades que de los mismos han de derivarse. En materia de mejora de la calidad de gestión pública, la ley se limita a destacar la conveniencia de incorporar a las políticas las mejoras recomendadas, pero sin establecer ningún procedimiento institucional concreto para promover este aprovechamiento. En este sentido, las diferencias con los sistemas considerados como mejores prácticas internacionales expuestos en la sección cuarta del trabajo son bastante evidentes.

Sobre la base del diagnóstico crítico ofrecido, el trabajo incluye una serie de consideraciones en relación con las carencias detectadas en el modelo institucional de evaluación pública propugnado por la Ley 27/2022. En él se destacan los puntos débiles y las posibles consecuencias de ese inadecuado tratamiento. El trabajo concluye con un esbozo de reforma, consistente en un listado de los elementos principales que, a nuestro juicio, deberían incorporarse para la construcción de un marco institucional de evaluación de políticas públicas que fuese realmente efectivo.

1. Introducción

Este documento tiene como objetivo abordar diferentes facetas del tema central de la calidad de las finanzas públicas, poniendo el foco principal en la evaluación de programas y políticas públicas. Nuestro propósito es cubrir los diferentes aspectos que, en nuestra opinión, constituyen los elementos claves de la relación existente entre evaluación y calidad de las finanzas públicas. En concreto, trataremos las cuestiones relacionadas con la institucionalización de la evaluación, la utilización de las evaluaciones como instrumentos de selección de los proyectos y programas que conforman las políticas públicas bajo criterios de eficiencia económica y eficacia, así como las potencialidades de la evaluación como instrumento para mejorar la gestión pública. En buena medida, las valoraciones y propuestas presentadas en este documento recogen, de forma sintética, las que se contemplan en los artículos publicados en el volumen 103 de *Ekonomiaz-Revista Vasca de Economía*, "Calidad de las finanzas públicas: El papel de la evaluación", I/2023, que he tenido la oportunidad de coordinar.

En este documento abordamos el tema de la evaluación desde la perspectiva de la modernización de la gestión pública prestando una especial atención a la reforma y creación de algunas instituciones públicas, así como la implantación, con vocación de habitualidad, de las herramientas de trabajo que deben emplearse por esas instituciones que conforman lo que habitualmente conocemos como marco institucional de la evaluación. En el trasfondo está la necesidad, a nuestro juicio, de poner en marcha un proceso de reforma institucional, imprescindible para aspirar a que la acción pública consiga logros efectivos y permanentes a medio y largo plazo en relación con el bienestar general, económico y social.

Tradicionalmente, el debate que ha predominado en relación con la influencia en el bienestar de los ciudadanos de la intervención de los gobiernos en la actividad económica de los países ha sido el que ha girado en torno al tamaño ideal de los gobiernos, identificado este, desde una perspectiva presupuestaria, con el nivel alcanzado de gasto público con relación al PIB. No fue hasta bien avanzados los años ochenta del siglo XX, cuando la preocupación por la eficiencia y la efectividad de las políticas públicas fue abriéndose paso desde una perspectiva fundamentalmente académica. Bien es cierto que esto sucedía en un momento en el que la mayoría de los sectores públicos de los países desarrollados se habían situado en sus niveles máximos de gasto público y, paralelamente, en sus niveles de presión fiscal y déficit público. Hasta entonces el interés por la eficiencia y la efectividad del sector público,

y en general por la calidad de las finanzas públicas, era bastante secundario. En el ámbito de la Unión Europea, la discusión tuvo bastante apogeo a lo largo de la década de los años noventa, si bien es cierto que impulsada, en buena medida, por las exigencias de disciplina fiscal.

En concreto, a finales de los años noventa del pasado siglo, los problemas de sostenibilidad de las finanzas públicas y la necesidad de aplicar programas de consolidación fiscal en los Estados miembros de la Unión Europea que iban a acceder a la última fase de la Unión Económica y Monetaria hizo que la atención se dirigiese hacia la introducción de mejoras en el funcionamiento de los sectores públicos, con el propósito de conseguir ahorros presupuestarios que permitiesen reducir el crecimiento del gasto público. Entonces, las propuestas se ocuparon, predominantemente, de la búsqueda de alternativas para reestructurar los programas de gasto, en la revisión de sus objetivos, así como de los criterios de elegibilidad de sus destinatarios, además de ampliar la presencia del sector privado en la oferta de servicios e infraestructuras de provisión pública, incluidas fórmulas de financiación que aliviasen la presión sobre las necesidades financieras del sector público.

Otro rasgo de esta etapa fue la preocupación por la relación existente entre la eficiencia del sector público y los aspectos organizativos e institucionales. En este sentido, se asumió con bastante aceptación –aunque con menor éxito real de implantación– que contar con una arquitectura organizativo-institucional adecuada, capaz de dar respuestas efectivas a los problemas de coordinación e incentivos propios de cualquier organización económica compleja, constituía un requisito esencial para alcanzar niveles satisfactorios de eficiencia en la gestión de los recursos públicos (Onrubia, 2005a).

Tras la crisis financiera de 2009, una nueva aproximación, la calidad de las finanzas públicas, se ha ido abriendo camino, incorporando algunas perspectivas novedosas a la hora de valorar en términos económicos las actuaciones públicas, especialmente las financiadas con recursos presupuestarios, aunque no exclusivamente.² Estamos ante, a nuestro juicio, un concepto con connotaciones valorativas atractivas, aunque hay que reconocer que los atributos de distinta índole subyacentes al término calidad hacen compleja su incorporación directa al análisis económico aplicado.

² El concepto de “calidad de las finanzas públicas” se asoció inicialmente con el diseño y aplicación de la política fiscal y fue incorporado como un nuevo objetivo estratégico de la Unión Europea en la Estrategia de Lisboa acordada en el Consejo Europeo celebrado en Lisboa el 23 y 23 de marzo de 2000, dentro de las acciones en materia de coordinación de las políticas macroeconómicas (saneamiento presupuestario, calidad y sostenibilidad de las finanzas públicas).

Desde una concepción inicial vinculada al desempeño de la política fiscal, el concepto de calidad de las finanzas públicas, como advierte Thöne (2023), abarca bastantes cuestiones cualitativas y estructurales, centrándose principalmente en el impacto que la composición del gasto público tiene a largo plazo en objetivos como el crecimiento económico, la sostenibilidad medioambiental, la inversión en infraestructuras y en capital humano, considerando como pilares fundamentales de esta última la educación, la sanidad, la política familiar y de género. Tras los exigentes procesos de consolidación presupuestaria puestos en marcha a raíz de la crisis financiera y de deuda pública, cabe hablar, como señala Rodríguez-Vives (2019), de la existencia de un interés renovado por el concepto de la calidad de las finanzas públicas, reconociendo su carácter multidimensional. Esto hace que su tratamiento en el ámbito de la evaluación sea abordable desde distintos ángulos, aunque desde una perspectiva aplicada se defiende que su meta, en última instancia, debería ser la búsqueda de una composición presupuestaria compatible con el crecimiento económico a largo plazo y la reducción de la desigualdad.

La experiencia pone de manifiesto que la calidad de las finanzas públicas va más allá de la determinación del volumen de recursos presupuestarios que deben ser asignados a los diferentes programas que articulan las políticas públicas, a la composición cualitativa del gasto público y a la forma concreta de su financiación. La calidad de las finanzas públicas, en última instancia, debe ponerse en relación con la capacidad que tienen los gobiernos, en sus diferentes niveles jurisdiccionales, para responder con efectividad a las demandas de la sociedad, tanto en lo concerniente a la provisión de bienes y servicios públicos como a la lucha contra la desigualdad y la pobreza. En definitiva, la calidad de las finanzas públicas es una condición imprescindible para promover el bienestar social a través del crecimiento económico y la redistribución de la renta y la riqueza.

Para que las políticas fiscales puedan ser consideradas de alta calidad, como señalan Afonso *et al.* (2005), deben fomentar el crecimiento económico, cumpliendo con los siguientes requisitos: a) proporcionar un entorno institucional favorable al crecimiento y unas finanzas públicas saneadas; b) limitar los compromisos del gobierno en la provisión de bienes y servicios a su papel esencial; c) establecer incentivos que promuevan el crecimiento para el sector privado y fomenten el uso eficiente de los recursos públicos; d) financiar las actividades del gobierno y, en su caso, las del sector privado, con un sistema fiscal eficiente y estable; e) apoyar la estabilidad macroeconómica mediante unas cuentas públicas estables y sostenibles.

La vinculación de la calidad con el desempeño del sector público en sus funciones requiere desarrollar una “buena gestión pública”. Esto supone considerar todos los elementos institucionales, organizativos, gerenciales y evaluativos que van a influir en el diseño y aplicación de las políticas públicas. No es difícil ver que este alcance “gerencial” de la noción de calidad va más allá del ámbito tradicional macroeconómico de la política fiscal.³ Pensamos que resulta difícil imaginar escenarios de intervención con unas finanzas públicas de calidad sin una identificación correcta de los objetivos a cubrir por las políticas públicas de acuerdo con las preferencias de la sociedad. Lo mismo sucedería sin la utilización sistemática de métodos basados en criterios económicos para la selección de proyectos alternativos y el diseño de los programas públicos. Igualmente, no sería concebible considerar unas finanzas públicas de calidad sin la realización regular de análisis de eficiencia y eficacia de esos proyectos y programas, sin la evaluación del impacto de las medidas adoptadas o sin la integración de los resultados de todos estos análisis y evaluaciones en los procesos de presupuestación.

En la relación que acabamos de exponer, la evaluación económica, en sentido amplio, desempeña un papel fundamental, sin el que estamos convencidos de que no es posible conseguir un grado satisfactorio de calidad en las finanzas públicas. En la práctica, la conexión de la calidad de las finanzas públicas con la gestión eficiente de los ingresos y gastos públicos depende, en gran medida, de la existencia de un sistema institucional de evaluación que funcione adecuadamente. En el ámbito que nos ocupa, la evaluación debe concebirse como un proceso generador de información que mediante el uso de herramientas diversas de análisis económico permite poner en relación los recursos destinados a la financiación de los programas y servicios públicos con los resultados, a corto, medio y largo plazo, alcanzados por estos. El propósito de esta generación de información no es otro que determinar, mediante técnicas cuantitativas, el grado de efectividad y de eficiencia, en el sentido de rentabilidad social, de esos programas y proyectos y, por agregación, de las políticas en las que se integran (Onrubia y Santín, 2017).

De este modo, la evaluación, en su concepción más amplia, constituye un instrumento imprescindible para generar la información necesaria para seleccionar los objetivos preferentes de las intervenciones públicas y asignar a cada una, bajo un criterio de eficiencia, los recursos presupuestarios necesarios para alcanzarlos, así como determinar la forma más neutral y equitativa de obtenerlos. Así, esta información se convierte en un elemento imprescindible para que las decisiones que

³ Sobre los principios de una buena gestión pública y su consecución, véase Albi y Onrubia (2018).

tienen que adoptarse en el desarrollo de la gestión pública sean consistentes con el uso eficiente de los recursos públicos. Solamente de esa manera, el funcionamiento de las unidades que conforman el sector público podrá desarrollar un comportamiento organizativo eficiente.

Además, debemos destacar otra valiosa función de la evaluación económica de las actividades públicas, tal y como la acabamos de describir. Nos estamos refiriendo a su capacidad para promover la transparencia en la actuación financiera de los poderes públicos y la rendición de cuentas en la gestión de recursos puestos a su disposición. Al someter la gestión pública a un escrutinio objetivo mediante la obtención de evidencia empírica se fomenta la apertura de la acción pública a la sociedad, reduciéndose los riesgos de despilfarro de los recursos, corrupción y de malversación de fondos.

En el volumen 103 de *Ekonomiaz* se han presentado una decena de artículos que abordan el tema central de la evaluación y la calidad de las finanzas públicas desde todas estas perspectivas apuntadas, y de los que podemos extraer interesantes consideraciones y recomendaciones para conseguir una mejor gestión de las finanzas públicas. Los artículos seleccionados cubren los diferentes aspectos que, como se acaba de exponer, representan los elementos claves del tema. A saber, las cuestiones relacionadas con la institucionalización de la evaluación, la utilización de las evaluaciones como instrumentos de selección, bajo criterios de eficiencia económica, de los proyectos y programas que conforman las políticas públicas, la integración de sus resultados en los procesos de presupuestación como elemento fundamental para el análisis, seguimiento y valoración de la efectividad de los programas de gasto o, desde una perspectiva macroeconómica, como instrumento para validar la calidad de la política fiscal.

En la sección final del trabajo presentamos una serie de consideraciones que, a nuestro juicio, podrían contribuir a dotar de contenido un proceso de reforma de la gestión pública basado en la implantación efectiva de un sistema integral de evaluación que funcione de forma sistemática y regular a lo largo de todas las fases de la intervención pública.

Tras esta introducción, el resto del trabajo se estructura de la siguiente forma. En la segunda sección se exponen los fundamentos de la evaluación de políticas públicas, con atención a su papel en el fomento de la calidad de las finanzas públicas. La sección tercera revisa, de forma sintética, las principales herramientas técnicas

disponibles para el desarrollo de las actividades de evaluación económica en sus diferentes categorías. El marco institucional de la evaluación pública es tratado en la sección cuarta, presentándose en primer lugar los aspectos teóricos que deben estar presentes en el diseño de un sistema institucional de evaluación; en segundo lugar, se presenta, de forma resumida, un panorama de las mejores prácticas internacionales en materia de evaluación pública. El bloque final del trabajo está dedicado a la evaluación pública en España. La sección quinta se encarga de presentar, de forma valorativa, la situación actual de la misma en el Estado y en las comunidades autónomas, desde sus antecedentes, fijados en la Constitución Española de 1978, hasta la reciente Ley 27/2022, de institucionalización de la evaluación de políticas públicas. El trabajo concluye con una sexta sección en la que se presentan las líneas que, a nuestro juicio, podrían esbozar una reforma de la evaluación pública en España que aspire a ser eficaz en sus metas.

2. La evaluación como instrumento de supervisión y fomento de la calidad de las finanzas públicas

Aunque con diferente intensidad según los países, las políticas de consolidación fiscal puestas en marcha por los gobiernos tras la crisis financiera y de la deuda pública supusieron importantes ajustes en los programas de gasto público, acompañados de aumentos impositivos y de los precios públicos. En España, las medidas afectaron prácticamente a todas las políticas de gasto público, tanto de competencia estatal como autonómica, sin olvidarnos de las de los gobiernos locales. Las políticas de gasto más afectadas fueron las relacionadas directa e indirectamente con la inversión pública, decisión explicada habitualmente por razones de cálculo político dado el menor coste electoral de estos ajustes a corto y medio plazo. Asimismo, el gasto correspondiente a la provisión en especie de bienes y servicios públicos, tanto en el caso de la producción pública como en el de la financiación de producción privada también experimentaron recortes significativos. En este caso, al no existir, por regla general, un reconocimiento legal de derechos económicos cuantificados a nivel individual, al contrario de lo que sucede con la mayoría de las prestaciones monetarias, la práctica de esos ajustes se veía facilitada, aunque inevitablemente con efectos sobre el acceso, amplitud y calidad de su oferta a los ciudadanos.

En escenarios de crisis económica, estos ajustes fiscales se adoptan con urgencia, por los rápidos deterioros que suelen experimentar los ingresos públicos. Esto hace que los ajustes consistan, en buena medida, en recortes lineales de los créditos presupuestarios inicialmente aprobados, si existen montantes sin ejecutar, o de disminución, cuando no supresión, de recursos destinados a programas de carácter plurianual. Además, la dificultad de modificar normativas legales que reconocen derechos individuales a percibir prestaciones con cuantías preestablecidas hace que una parte importante de los recortes en los programas se concentre en el gasto público correspondiente a la provisión de servicios públicos, como la educación, la sanidad o los servicios sociales (por ejemplo, no reponiendo vacantes entre los empleados públicos o congelando las transferencias a las empresas contratistas en los conciertos de provisión), ya que resulta más sencillo modificar o restringir las condiciones de acceso a los mismos, cuando no directamente suprimirlos. No obstante, algunas prestaciones monetarias, cuyas cuantías individuales no suelen establecerse por ley (como muchas becas, ayudas a la investigación o compensaciones de gastos por dependencia), también se ven afectadas por los ajustes. Sin embargo, es la inversión pública la categoría de gasto que suele soportar un mayor impacto en la consolidación fiscal, a pesar de su contrastado efecto sobre

el crecimiento económico. De hecho, en España, a pesar de la fuerte entrada de fondos europeos procedentes del programa Next Generation EU, la inversión pública en 2022 seguía situándose un 55% por debajo del máximo alcanzado en 2009 (Fundación BBVA e IVIE, 2023).

2.1. El papel de la evaluación de políticas públicas

Como hemos apuntado, el diseño de las estrategias de consolidación suele estar apremiado por la necesidad de reducir el rápido aumento del endeudamiento público al que conduce el mal comportamiento de la recaudación fiscal. Esto hace que los recortes se adopten sin tener en cuenta el grado de eficiencia concreto con el que se ejecuta cada programa de gasto, ni su efectividad medida en términos de impacto económico y social. Esta forma de proceder genera costes de bienestar nada despreciables. De hecho, como argumentan Onrubia y Sánchez-Fuentes (2017), una parte importante de esos costes de bienestar podrían evitarse o al menos mitigarse si las reducciones del gasto fuesen selectivas y se adoptasen a partir de la información proporcionada por la evaluación económica de los programas y proyectos públicos. Los ahorros que podría generar una evaluación rigurosa del gasto público convierten a esta en una herramienta con enorme valor para afrontar estrategias de consolidación fiscal, contribuyendo además a obtener nuevas ganancias de bienestar social en el medio y largo plazo.

De acuerdo con lo anterior, la utilización de los resultados de la evaluación en el diseño de estrategias de reestructuración de las políticas públicas proporcionaría una suerte de "doble dividendo" por adoptar la terminología habitual en las políticas medioambientales. Primero, porque permite reducir el gasto menos productivo en términos sociales, facilitando además la supresión de aquellas actividades y programas asentados en el presupuesto fundamentalmente por razones de cálculo político, sin que exista una justificación clara en términos de eficiencia o equidad (como suele suceder, por ejemplo, en muchos de los beneficios fiscales reconocidos en las normas tributarias o en determinadas subvenciones sectoriales). Y segundo, porque una fracción de los recursos presupuestarios liberados por los ajustes adoptados podrían reorientarse hacia proyectos y programas de gasto socialmente más rentables.

En el caso de España, las presiones del envejecimiento poblacional sobre importantes políticas de gasto, como las pensiones, la sanidad o la dependencia; la necesidad de una reforma integral de nuestro bastante mejorable sistema educativo en todos sus

niveles; la potenciación de una política eficaz en I+D+i y transformación digital empresarial, junto con los objetivos medioambientales derivados del cambio climático; o las exigencias en materia de seguridad y defensa que imponen los nuevos escenarios de conflicto y de terrorismo internacional, sitúan a nuestras finanzas públicas ante enormes retos que van a tensionar el gasto público en los años venideros. De hecho, como se ha visto en la reciente pandemia de la Covid-19, esto ya lleva sucediendo algunos años.

En consecuencia, abordar un plan de reforma racionalizadora del gasto público, en términos de eficiencia y efectividad, emerge en España como una meta de la agenda política estatal, autonómica y local, imprescindible e inaplazable. Un plan de este tipo resulta esencial para poder valorar la sostenibilidad a medio y largo plazo de las finanzas públicas y, en definitiva, del Estado del Bienestar. El plan contaría con varias etapas. En la primera, se debería comenzar por una revisión sistemática de todos los programas presupuestarios en vigor, que incluya: a) un análisis riguroso de la estructura y composición de gasto destinado a cada uno, considerando y valorando las alternativas de implementación disponibles; b) un análisis del grado de desempeño alcanzado por estos programas en sus años de funcionamiento; c) una valoración de sus resultados en términos de impacto económico y social, basada en indicadores objetivos y contrastables. En una segunda etapa, la revisión realizada en la primera debería dar lugar a un informe en el que se mostrase cómo ha cambiado la estructura de los programas de gasto a lo largo del tiempo, qué niveles de gobierno han puesto en marcha cambios de gestión y han practicado ajustes presupuestarios de manera motivada, valorando en todos ellos sus efectos en relación con los objetivos perseguidos.

La comparación de los procesos seguidos por los países desarrollados y, en especial los de la Unión Europea, considerados como de "mejores prácticas", tal y como se defiende en Albi y Onrubia (2016), debería servir para orientar las reformas del gasto público a poner en marcha. No obstante, también somos conscientes de que, si los cambios no se ven acompañados de medidas de modernización de la presupuestación que integren los resultados de la evaluación de programas y proyectos como las apuntadas por Baños-Rovira (2023) en su artículo del número 103 de *Ekonomiaz* o del marco de control y de transparencia del gasto propuestas por Albi y Onrubia (2018), el éxito de estas reformas quedaría seguramente comprometido de forma seria. Pero de todas las acciones a poner en marcha, la reforma de la evaluación de las políticas y programas de gasto, y fundamentalmente la de su institucionalización, es a nuestro juicio la principal de ellas, además de su interrelación con el resto de los

elementos de la gestión de las finanzas públicas. Cabe decir que sin una evaluación económica que funcione adecuadamente, conocer la eficiencia e impacto económico y social de los proyectos y programas públicos es una tarea bastante quimérica, como en algunas de las bien intencionadas propuestas de reforma iniciadas en España se ha podido comprobar desde los años noventa. En definitiva, como se ha dicho, los gobiernos responsables de las finanzas públicas se enfrentan aquí a una exigencia incuestionable para poder realizar reorientaciones efectivas del gasto público, sin las cuales a largo plazo será muy complicado dar respuestas satisfactorias a unas necesidades de la sociedad cambiantes por los retos señalados.

En esta línea, hay que valorar positivamente las actividades de evaluación del gasto público que viene desarrollando la Autoridad Independiente de Responsabilidad Fiscal (AIReF). Aunque inicialmente esta función de evaluación de políticas públicas no estaba contemplada en su estatuto orgánico⁴, AIReF la ha ido consolidando a lo largo del tiempo, tal que en la reestructuración adoptada por la institución en septiembre de 2021, dentro de su Plan Estratégico 2020-2026, incluyó en su estructura una nueva División de Evaluación del Gasto Público. Esta modificación organizativa fue adoptada en cumplimiento del Componente 29 (Mejora de la eficacia del gasto público) del Plan de Recuperación, Transformación y Resiliencia. El núcleo fundamental de esta actividad corresponde a la gestión y realización de las Spending Review (revisiones de gasto), proceso de evaluación del gasto encargado por el Gobierno central y que actualmente alcanza su tercer paquete de revisiones.⁵ Además, la AIReF realiza otras evaluaciones por encargo de gobiernos autonómicos.⁶ Por mandato legal específico, la AIReF lleva a cabo desde 2021 la evaluación de los

⁴ El artículo 3 del Real Decreto 215/2014, de 28 de marzo, por el que se aprueba el Estatuto Orgánico de la Autoridad Independiente de Responsabilidad Fiscal, establece que "1. La Autoridad Independiente de Responsabilidad Fiscal tiene como fin garantizar el cumplimiento efectivo por las Administraciones Públicas del principio de estabilidad presupuestaria consagrado en el artículo 135 de la Constitución Española, en los términos en que ha sido desarrollado en la Ley Orgánica 2/2012, de 27 de abril; 2. Para ello realizará una evaluación continua del ciclo presupuestario, del endeudamiento público y un análisis de las previsiones económicas".

⁵ El proyecto de análisis de gasto público conocido como Spending Review fue iniciado con el encargo realizado por el Gobierno de España como consecuencia de la Actualización del Programa de Estabilidad de 2017-2020. Una vez concluida, se inició la Spending Review II 2018-2021, que contempló tres fases y cuyos resultados se encuentran ya publicados. Actualmente, está en fase de desarrollo el encargo de la Spending Review III 2022-2026.

⁶ AIReF ha llevado a cabo, hasta la fecha, diversas evaluaciones para las comunidades autónomas de Andalucía (del sistema universitario público andaluz), Aragón (del gasto público farmacéutico, de la política presupuestaria y de personal en educación no universitaria y del modelo de financiación de la Universidad de Zaragoza), Baleares (revisión del gasto público en sanidad y educación), Cantabria (de las empresas y fundaciones públicas autonómicas), Castilla y León (de las políticas activas de empleo, de la institucionalización de la evaluación de políticas públicas y de la gestión de la publicidad institucional), Cataluña (de la valoración de la concesión para construcción del "Eix transversal C-25"), Extremadura (de las políticas activas de empleo y del gasto público farmacéutico y en equipos de alta tecnología) y Navarra (revisión del gasto público en sanidad).

resultados de la aplicación del programa estatal del Ingreso Mínimo Vital aprobado en 2020, así como del Fondo para la Promoción del Desarrollo, instrumento financiero principal del sistema español de cooperación internacional para el desarrollo. Anteriormente, en 2019, publicó un estudio en el que se evaluaban los distintos programas de rentas mínimas, competencia de las comunidades autónomas. En el ámbito local, la AIREF ha realizado estudios en ámbitos muy concretos.

Las evaluaciones de gasto incluidas en las sucesivas Spending Review tienen su origen en el compromiso contenido en el Plan Presupuestario remitido a la Comisión Europea en diciembre de 2016 y que se recogió en el Programa de Estabilidad 2017-2020. El propósito de estas Spending Review, como se recoge por la propia AIREF, es mejorar la eficiencia del gasto para el conjunto de las Administraciones Públicas⁷. Inicialmente la AIREF elaboró un Plan de Acción, aprobado en el Consejo de Ministros del 26 de enero de 2018, que concretaba la actuación en siete proyectos de evaluación que afectan tanto a Comunidades Autónomas como a la Administración Central. Este Plan de Acción fue diseñado de acuerdo con los principios comunes establecidos por el Eurogrupo para Spending Review bajo las premisas de un compromiso político firme y duradero, el diseño e implementación, seguimiento y comunicación, y coherencia en la planificación presupuestaria. Como expresan Gambau y López-Herrera (2023) en su artículo incluido en el volumen 103 de *Ekonomiaz* en relación con el encaje que las Spending Review pueden tener en el reciente marco normativo aprobado en España para la institucionalización de la evaluación de políticas públicas en la Administración General del Estado, aunque el balance puede considerarse positivo, es posible identificar algunos retos y riesgos que no deben dejarse de lado en el debate público y político para consolidar de forma efectiva la institucionalización de la práctica evaluadora en su concepto más integral.

El análisis de las Spending Review realizadas permite a estos autores afirmar que, aunque con distintos alcances y sistemáticas de trabajo en función de la materia, todas ellas han apuntado recomendaciones interesantes con propuestas de mejora que, en mayor o menor medida, se han tratado de poner en marcha por los organismos responsables de gestionar los programas de gasto analizados. No obstante, advierten de la indudable orientación ex post de estas Spending Review, algo que no responde al concepto de evaluación integral por el que se apuesta como objetivo en el Plan de Recuperación, Transformación y Resiliencia, si bien valoran la

⁷ La información sobre la Spending Review puede consultarse en su web institucional <http://www.airef.es/spending-review>

experiencia acumulada que debería ser capitalizada por el nuevo marco normativo e institucional.

2.2. **Ámbito funcional de la evaluación pública**

En relación con el alcance de la evaluación de políticas, programas y proyectos públicos, es necesario delimitar previamente su ámbito funcional. Así, debemos destacar la importancia que deben tener el enfoque económico y el análisis cuantitativo dentro de la evaluación. Aunque la transversalidad es un rasgo distintivo presente en las actuaciones públicas, explicado por su amplia diversidad sectorial, evaluar los resultados alcanzados por las mismas, la eficiencia con la que se utilizan los recursos o el impacto sobre el bienestar social, es una actividad que requiere una medición cuantitativa, sujeta a estándares rigurosos y aceptados de forma unánime por los especialistas y la literatura. De lo contrario, la evaluación termina convirtiéndose en simples ejercicios de retórica, con escasa utilidad para influir en la calidad de la gestión pública y para fortalecer la exigencia de responsabilidades vinculada a la rendición de cuentas de los políticos y gestores públicos. Parece poco discutible que los resultados de las evaluaciones pueden y deben ser valorados desde diferentes enfoques, con el ánimo de mejorar el diseño de las políticas desde el punto de vista social. Sin embargo, se defiende aquí que esto no debe suponer una renuncia a las técnicas y métodos disponibles en el análisis económico cuantitativo, instrumentos indispensables para obtener valoraciones contrastables del buen uso de los recursos públicos.

A lo anterior hay que añadir que el uso de datos, métodos estadísticos y análisis económico en la medición de costes y en la evaluación de los efectos de las políticas públicas constituye una herramienta fundamental para quienes deben diseñarlas y aplicarlas. Como se ha dicho, el análisis del gasto público y de las propuestas normativas, realizado con independencia del poder político, es un elemento crítico en los procesos de reforma del gasto público y en la configuración de estrategias de consolidación fiscal. Por supuesto, la necesidad de incorporar la evaluación de forma regular y sistemática requiere contemplar los condicionantes del diseño institucional, sin los cuales su aplicación termina siendo escasamente efectiva.

Como se destaca en Albi y Onrubia (2018), para que contribuyan a mejorar el bienestar social, las decisiones colectivas deben ser adoptadas por los poderes públicos con racionalidad económica e imparcialidad social. Esto exige que los decisores políticos, que han de sopesar puntos de vista diferentes sujetos a muy

diversas presiones, actúen con criterios de eficiencia económica y, por supuesto, también de equidad. Por tanto, la calidad de las finanzas públicas debe ser entendida en el proceso de esta consecución del bienestar social, lo que convierte a los métodos de evaluación en instrumentos indispensables para alcanzar ambos.

La evaluación de políticas públicas puede definirse como un “proceso generador de información e indicadores relevantes que permite, a través de distintas herramientas de análisis económico, relacionar los recursos, las actividades y los resultados a corto, medio y largo plazo de los programas y servicios públicos para cuantificar y determinar si los mismos resultan eficientes o socialmente rentables. El objetivo de la evaluación es utilizar la información que se deriva del análisis en el apoyo de una toma de decisiones que permita mejorar el nivel de eficacia, eficiencia y/o la equidad del programa o servicio público analizado” (Albi y Onrubia, 2015:114). Su función consiste en dotar a los gobiernos e instituciones públicas de instrumentos de ayuda que mejoren la adopción de las decisiones colectivas proporcionándoles, al mismo tiempo, mayor control y transparencia.

Los resultados de la evaluación ex post de las actuaciones públicas, en un marco de transparencia, constituyen un elemento indispensable para un adecuado funcionamiento de la rendición de cuentas y la exigencia de responsabilidades sobre la gestión realizada. Sin esta información proporcionada por la evaluación económica, resulta imposible ir mucho más allá de la noción tradicional de control vinculada al cumplimiento de los principios de legalidad y regularidad contable. Hay que aclarar que los instrumentos de análisis empírico empleados en la evaluación tienen sus limitaciones; y su aplicación puede, en muchos casos, no ofrecer conclusiones definitivas, ni tan siquiera en términos de eficiencia económica, y mucho menos en los de equidad, como es sabido, afectados por juicios de valor. Pero esto no obsta para que sean imprescindibles para una gestión pública de calidad, si consideramos su capacidad para generar información de utilidad tanto para la toma de decisiones como para la valoración de los resultados alcanzados y, consecuentemente, hacer realidad la exigencia de responsabilidades en la gestión de los recursos públicos.⁸

Es necesario advertir que para que estos instrumentos sean realmente útiles, es fundamental disponer de información de calidad utilizable por los métodos de evaluación empleados, de manera que permitan su adecuado tratamiento estadístico y económico. En este sentido, el *Big Data* abre unos horizontes impredecibles para

⁸ Véase Mayne (2017) sobre la relación entre la práctica institucional de la evaluación, la exigencia de responsabilidades y la mejora de la gestión pública.

la evaluación de políticas y programas públicos, lo que es esperable que favorezca mejores asignaciones del gasto público y, en definitiva, una mayor calidad de la gestión pública⁹. Como destacan Finucane *et al.* (2018: 121), “en esta era del *Big Data*, la demanda de evaluaciones de programas que son simultáneamente menos costosas y más informativas va en aumento, algo que presiona a los investigadores y analistas a aprovechar al máximo el volumen y la velocidad con la que los datos se recopilan rutinariamente durante la investigación de políticas sociales, con el fin de responder de manera eficiente a la importante pregunta de qué funciona para quién”. No obstante, advierten que los investigadores podrían verse tentados a sacrificar el uso de métodos rigurosos en la evaluación del impacto de los programas públicos, en beneficio de la urgencia de los resultados, algo que podría restar validez a los análisis realizados. Estos autores recomiendan el uso de técnicas de inferencia causal que aprovechen esos inmensos volúmenes de datos, en tiempos razonables, pero sin renunciar a los principios metodológicos que aseguran la consistencia de los análisis realizados.

En relación con la aplicación de los métodos de evaluación, no puede olvidarse que se trata de elegir entre una diversidad de técnicas, por lo general, bastante complejas, sometidas a constante actualización. Acertar en la elección es fundamental para la validación de los resultados de las evaluaciones. Y esto requiere un análisis previo, de acuerdo con la naturaleza y contenidos de los programas y proyectos a evaluar. No es lo mismo evaluar una reforma laboral que una reforma fiscal, que medir los resultados de un protocolo terapéutico a aplicar en hospitales públicos, que un programa de acceso a viviendas sociales, que un programa de garantías de rentas mínimas como el Ingreso Mínimo Vital, o que un paquete de medidas de protección medioambiental. Asimismo, el ámbito de aplicación de los programas y proyectos públicos también resulta bastante relevante, especialmente en lo concerniente a la interrelación de los programas evaluados con otras políticas o con las regulaciones normativas que suelen acompañarlos. Como ejemplos de estos condicionantes de la evaluación, podríamos considerar dos de ellos: lo que supone para un proyecto concreto de transporte de viajeros mediante una línea de alta velocidad ferroviaria considerar los impactos económicos que exceden a sus usuarios directos en zonas geográficas muy amplias, o lo que implica, por ejemplo, para políticas de protección del medioambiente. Y lo que implica para la evaluación de un programa de prestaciones de desempleo el tratamiento de las rentas salariales en el IRPF o la existencia de otras ayudas sociales. Además, hay que tener en cuenta que,

⁹ Una interesante revisión sobre las potencialidades del *Big Data* en la gobernanza pública, incluida la evaluación de políticas públicas, puede verse en Misuraca *et al.* (2014).

en una evaluación integral de un programa de gasto público también es necesario valorar el coste de bienestar que supone la aplicación de los impuestos con los que se obtienen los recursos necesarios para su financiación, en términos del concepto del coste marginal de los fondos públicos.¹⁰

¹⁰ En relación con el concepto de coste marginal de los fondos públicos y su aplicación a la evaluación, véase Dahlby (2008).

3. Herramientas al servicio de la evaluación económica de las políticas públicas

En esta sección ofrecemos una revisión del papel de la evaluación económica desde una perspectiva instrumental. Para ello, agrupamos las técnicas más habitualmente empleadas diferenciando tres categorías principales de la evaluación, cuyos contenidos abordaremos seguidamente en este orden: a) los métodos de selección de proyectos de gasto público, integrantes fundamentales de la evaluación ex ante; b) las técnicas de medición de la eficiencia orientadas tanto a la gestión de los programas y proyectos de gasto público como al funcionamiento de las unidades organizativas que los implementan; c) las técnicas utilizadas para evaluar el impacto en el bienestar de los ciudadanos de esos programas y proyectos, así como las normativas reguladoras puestas en marcha por los gobiernos.

En el primer bloque aparecen las dos técnicas tradicionalmente empleadas en la evaluación prospectiva de proyectos públicos: el Análisis Coste-Beneficio (ACB) y su simplificación en forma de Análisis Coste-Efectividad (ACE). Como se argumenta en Albi y Onrubia (2018), ambas trascienden a lo que son, en sentido estricto, técnicas de evaluación. Estaríamos ante lo que puede denominarse un “enfoque coste-beneficio”, en el sentido de paradigma de evaluación general de proyectos y programas. Dicho de otra manera, estaríamos ante una forma de pensar en relación con la valoración económica de las políticas públicas. El segundo de los bloques que identificamos incluye los métodos de medición de la eficiencia productiva, empleados como técnicas dirigidas a medir el comportamiento de las unidades institucionales encargadas de la producción de bienes y servicios públicos o de financiar su provisión. Este tipo de análisis se realiza desde una perspectiva de eficiencia operativa, tanto en su vertiente técnica –la más habitual– como asignativa. Por último, el tercer bloque está formado por los métodos utilizados para realizar evaluaciones económicas del impacto de programas, servicios públicos o regulaciones normativas, con el propósito de valorar en qué medida su aplicación alcanza los objetivos perseguidos. No debe pasarse por alto que el logro de esos objetivos sirve de justificación para su adopción, con la correspondiente asignación de recursos a través de los presupuestos públicos.

La consideración de la evaluación como instrumento favorecedor del uso eficaz y eficiente de los recursos públicos requiere contemplar tanto la vertiente ex ante como ex post de esta (Albi y Onrubia, 2016). En la primera, las técnicas empleadas tienen por objeto evaluar, de forma comparativa, las alternativas de las que se dispone para

desarrollar los programas o proyectos concretos, con el fin de elegir la más idónea, atendiendo a un criterio de rentabilidad social. Esto supone ordenar las alternativas disponibles a partir de cuantificaciones monetarias de las potenciales ganancias de bienestar social. Las técnicas del ACB y el ACE son los instrumentos habituales para realizar las evaluaciones ex ante, aunque a veces ambas también son empleadas como métodos de evaluación ex post. Por su parte, las técnicas de medición de la eficiencia productiva y las de evaluación de impacto constituyen los métodos aplicables en la evaluación ex post. Su utilización tiene una vocación correctora respecto del funcionamiento de las políticas, si bien las técnicas de evaluación de impacto también se emplean a veces con carácter prospectivo en el diseño de proyectos y programas¹¹. En un adecuado funcionamiento del sistema institucional de evaluación de la actividad del sector público, ambas vertientes temporales deben estar configuradas de manera complementaria.

3.1. El Análisis Coste-Beneficio (ACB)

El ACB es la técnica más completa para valorar la relación entre los recursos empleados (costes), los objetivos perseguidos y los resultados alcanzados (beneficios). Este método consiste en valorar las corrientes temporales de todos los beneficios y costes sociales de cada una de las alternativas disponibles de desarrollo de un proyecto o programa. Esos flujos de gastos e ingresos, definidos en términos sociales, se transforman en Valor Actual Neto (VAN) mediante la aplicación de una tasa de descuento, también definida en términos sociales. En el ACB, el VAN es utilizado como regla de decisión para elegir la alternativa que se estima que debe proporcionar el mayor valor esperado. Otra forma de seleccionar las alternativas consideradas mediante el ACB consiste en ordenar la rentabilidad social de estas, a través del cálculo de las respectivas Tasas Internas de Rentabilidad (TIR), definidas también en términos sociales.

Desde su aparición en el ámbito de la literatura económica en la década de los años sesenta, el ACB se ha considerado adecuado para la evaluación de proyectos de inversión pública como los de infraestructuras o servicios públicos con un alto componente de inversión, por ejemplo, los relacionados con el transporte público.¹² No obstante, también se aplica a programas de las principales áreas de la política pública como la educación, la sanidad o el medioambiente. Es una técnica con mucha

¹¹ Sobre el uso prospectivo y retrospectivo de las técnicas de evaluación de impacto, puede verse el trabajo de Gertler *et al.* (2011).

¹² Una revisión de la literatura seminal sobre ACB puede verse en Prest y Turvey (1966).

potencialidad, pero cuyas dificultades de aplicación en proyectos públicos concretos no son pequeñas.¹³

Su aplicación exige considerar la totalidad de los beneficios y costes sociales de un proyecto o programa, tanto financieros como no financieros, tangibles e intangibles, además de la valoración monetaria de los efectos externos, positivos y negativos, evitando una "doble contabilización". Los costes y los beneficios tienen que considerarse, como hemos apuntado, desde un punto de vista social, algo que complica su valoración, en comparación con la aplicación de las técnicas del VAN y el TIR en la evaluación de proyectos privados. Esa valoración social consiste en incorporar al análisis los incrementos y disminuciones que se espera que se produzcan en los excedentes de los consumidores y los productores afectados por el proyecto, algo que no siempre se refleja en términos estrictamente financieros.

Además, debe tenerse en cuenta la existencia de efectos externos, positivos y negativos, así como calcular "precios sombra" que establezcan la valoración social de las partidas de costes y beneficios en términos marginales. Otra cuestión complicada en la utilización del ACB es la elección de una tasa de descuento social. La elección de la regla de decisión que permita ordenar las alternativas también es otro elemento de discusión abierta en el uso del ACB. Asimismo, las dificultades de aplicación aumentan cuando los proyectos o programas evaluados se encuentran afectados de forma significativa por escenarios de riesgo o incertidumbre.

A la hora de identificar a los beneficiarios y perjudicados del proyecto o programa evaluado, hay que tener en cuenta que los beneficios de una actuación pública pueden ser generales y aparecer dispersos entre toda la población. En el caso de los costes, estos se deben considerar en términos de coste de oportunidad, es decir, recogiendo adecuadamente el valor de la mejor oportunidad perdida por usar los recursos de forma alternativa, lo cual puede afectar también de forma general a todos los individuos o grupos implicados.

De forma intuitiva, el ACB es un método que tiene como finalidad valorar los efectos positivos y negativos ocasionados por un cambio de política o acción, en términos sociales, para que se pueda decidir si la sociedad prefiere ese cambio, otro o ninguno. La utilización del concepto económico de excedente como medida monetaria del cambio de bienestar permite cuantificar las alteraciones que esos efectos producen

¹³ El detalle de esta técnica, sus desarrollos metodológicos y sus campos concretos de aplicación pueden consultarse en Albi (1989), Brent (2006) y De Rus (2021).

en el bienestar de cada individuo o, de forma agregada, en los grupos de individuos afectados por el proyecto público a través de la suma o resta (con sus signos correspondientes) de los excedentes de todos los agentes económicos afectados por la actuación pública evaluada. Esa suma puede ser ponderada si se quiere incorporar un criterio de equidad a la evaluación. Si la suma es positiva, significa que el proyecto es deseable socialmente.

Hay que precisar que el marco de evaluación a través del uso del ACB que acabamos de exponer refleja un escenario de análisis de equilibrio parcial. Tal es así que las funciones de demanda solo dependen del precio, pero no tienen en cuenta la renta real de los individuos o las interacciones con otras políticas o con otros mercados. Además, se parte del supuesto de que la variación de precios ocasionada por un proyecto o programa solamente origina un efecto renta, siendo el efecto sustitución nulo o muy pequeño. Esto supone admitir que, en este enfoque de equilibrio parcial, las variaciones de precios no se consideran importantes y no alteran la renta real, lo que equivale a suponer que no existen efectos distributivos o que estos son poco relevantes. Obviamente, este supuesto no sería aceptable con variaciones importantes de precios, lo que obligaría a utilizar funciones de demanda y oferta compensadas y a utilizar los conceptos de variación compensatoria o equivalente, en vez de los excedentes del consumidor y del productor, lo que hace más complejo el análisis.

Otro problema a tener en cuenta en la aplicación del ACB es el que afecta al momento de valoración del proyecto público, según consideremos una estructura de precios relativos previa o posterior a su entrada en funcionamiento. De este modo, tener en cuenta una u otra estructura de precios relativos puede conducir a que se acepte o rechace el mismo proyecto. Desde la óptica de la compensación expuesta, sería posible que los beneficiarios pudieran compensar a los perjudicados para que aceptaran el cambio social que supone el proyecto y, a su vez, que estos pudieran compensar a los primeros para volver al *statu quo* original con la estructura de precios inicial. Si esto sucede, el ACB no resulta un criterio de decisión social totalmente fiable y mucho menos infalible. Si bien es cierto que la literatura especializada ha avanzado en la propuesta de criterios más rigurosos, es necesario advertir de este tipo de cuestiones y debilidades metodológicas que pueden surgir en la utilización del ACB.

En resumen, como expresa Albi (1989), el ACB no puede considerarse, sin más, un método de decisión social concluyente para elegir uno entre varios proyectos

alternativos. Sin embargo, como señaló atinadamente Williams (1972), uno de los precursores del ACB, si se acepta el principio de compensación potencial y se separan las cuestiones de eficiencia y de equidad en la evaluación, la suma algebraica de los efectos positivos y negativos de un proyecto, o de una política pública, brinda un método de análisis bastante razonable para poder descartar aquellas alternativas que nunca debieran aplicarse. Pero, como hemos visto, el ACB es complejo de utilizar, lo que recomienda una aplicación flexible de su metodología, adaptada en cada caso a las características propias de los proyectos o programas considerados. Por último, debe advertirse que una aplicación correcta del ACB requiere contar con grupos de expertos en evaluación con buena formación en Economía Pública y en Economía del Bienestar, así como con el apoyo de especialistas sectoriales según el ámbito técnico que afecta a cada proyecto. Por todo ello, como advierten Albi y Onrubia (2018), el ACB no es un tipo de evaluación que deba institucionalizarse para su aplicación de forma generalizada y mecánica, y menos para su uso detallado y profundo.

Posiblemente, la mayor utilidad del enfoque de evaluación "coste-beneficio" es contar con una metodología normalizada, de manera que se puedan realizar, preferiblemente por una institución competente, evaluaciones simplificadas de los efectos positivos y negativos de los proyectos y programas cuantitativamente más relevantes que conforman las políticas públicas, contemplando sus costes y beneficios de forma conjunta y atendiendo a una perspectiva de los efectos generados sobre los agentes económicos, lo más amplia posible, que podemos identificar como "valoración en términos sociales". Por supuesto, no se trata tampoco de seguir una metodología de análisis simplificada sin más: es necesario conocer y establecer, de forma explícita, los supuestos en los que se basan esas simplificaciones. También hay que tener en cuenta que el ACB está construido fundamentalmente para evaluar en términos de eficiencia, de forma que los cambios distributivos generados no suelen ser contemplados o, si se consideran, solamente es de forma residual. Por ello, es recomendable en su aplicación acompañar la realización de una evaluación ACB con un análisis complementario de los efectos distributivos generados por el efecto, en línea con lo propuesto por Onrubia y Rodado (2023) en su artículo sobre la evaluación de los aspectos redistributivos de las políticas públicas incluido en el número 103 de *Ekonomiaz*.

3.2. El Análisis Coste-Efectividad (ACE)

El ACE, también denominado análisis "coste-eficacia", mantiene en lo sustancial el enfoque del ACB, aunque se trata de una variante más simplificada. En concreto, el

ACE se suele aplicar cuando, siendo posible calcular los costes sociales de un proyecto, resulta difícil disponer de buenas estimaciones monetarias de sus beneficios. En ese caso, el análisis a través de ACE se desarrolla mediante la comparación de los costes con índices o medidas no monetarias de la efectividad de los proyectos o programas. Se trata de contar con una medición de la eficacia del gasto que supone el proyecto de acuerdo con sus objetivos (por ejemplo, el coste que aparea la reducción de los índices de mortalidad o morbilidad, de los de criminalidad o la reducción de los accidentes de tráfico, o el aumento del rendimiento académico de los titulados universitarios o del grado de satisfacción con un servicio). De acuerdo con este contenido, la interpretación de los resultados del ACE es bastante intuitiva, lo que lo convierte en una técnica atractiva para los decisores públicos: de las opciones consideradas para desarrollar una política, ¿qué alternativa supondrá un coste menor cuando se trata de alcanzar un nivel de eficacia determinado? Esta eficacia o efectividad se mide a través de uno o varios indicadores, adecuadamente relacionados con los objetivos perseguidos. De otra forma, se trata de ver qué alternativa permitirá alcanzar el mayor nivel de efectividad para un determinado coste, o el programa que maximiza la efectividad por unidad de coste (teniendo en cuenta que pueden existir restricciones respecto del nivel absoluto de efectividad o de coste).

La economía de la salud es un ámbito de aplicación idóneo del ACE, en el que con frecuencia se emplea la variante denominada "análisis coste-utilidad". Esta técnica suele aplicarse en la selección de tratamientos médicos alternativos susceptibles de ser valorados de acuerdo con el número de años en los que se alargaría la vida de los pacientes que los reciben y además son objeto de valoración. Para ello, una medida adecuada de efectividad en asuntos relacionados con la esperanza de vida es considerar los "años de vida ganados ajustados por su calidad" (QALY por sus siglas en inglés) (Williams *et al.*, 1987), o los denominados "índices de envejecimiento saludable" como el propuesto por Wachs *et al.* (2020).

La evaluación de proyectos de gasto social es otro ámbito natural de utilización del ACE. Esto se explica porque el ACE prioriza encontrar indicadores de resultados, desempeño o efectos del proyecto relevantes, sencillos y, por lo general, menos costosos de conseguir; a diferencia del ACB, que valora monetariamente los beneficios de la actuación pública. No obstante, para que la técnica ACE sea útil es necesario que los indicadores de eficacia utilizados estén fuertemente relacionados con el objetivo u objetivos finalistas de los proyectos o programas evaluados. En el caso de que existan varios objetivos, resulta necesario ponderarlos, integrándolos en

un único indicador que represente el grado de consecución de un objetivo sintético. En estos casos, el análisis multicriterio puede ser también un buen instrumento para aplicar el ACE.

A pesar de las ventajas señaladas, el ACE es un método no carente de críticas. Las principales se centran en las dificultades para establecer relaciones consistentes entre costes y efectividad. Una forma de solventarlo es realizar trabajos empíricos previos para contrastar la robustez de esas relaciones, que sirvan para validar la idoneidad de los indicadores finalmente seleccionados.

Otro requisito respecto de la definición de los indicadores en el ACE es evitar la confusión entre *inputs* con *outputs* y, respecto de estos, tratar de no sustituir indicadores de resultados, desempeño o actividad (*performance*) por indicadores de carga de trabajo o de tareas intermedias, como el número de alumnos de una escuela o de pacientes atendidos en un centro sanitario. También debe evitarse la utilización de medidas de "costes promedio", por ejemplo, los costes por alumno o por paciente, ni ratios de alumnos por profesor o de médicos por cama de un hospital. Ninguna de estas medidas aporta información sobre la efectividad del gasto empleado en los programas. Los indicadores utilizados deben estar correlacionados con los objetivos perseguidos por la actuación pública, de tal manera que los indicadores los hagan explícitos. Por el contrario, son indicadores adecuados, por ejemplo, los tiempos de curación de un tratamiento, la tasa de abandono en ciclos educativos, el porcentaje de reducción de accidentes de tráfico o un indicador de la gravedad de las lesiones. Debe advertirse que elegir malos indicadores puede introducir sesgos indeseables en la gestión de los servicios vinculados al programa.

En todo caso, se debe partir de la premisa de que ningún indicador de los resultados o de la efectividad de un programa o proyecto va a ser perfecto, ni va a poder medir con precisión el valor económico del *output* público. Pero hay que asumir que un indicador imperfecto de los resultados de un programa seguramente será mejor que ninguna información, si bien es cierto que los indicadores siempre deben representar medidas razonables de los objetivos perseguidos. En este sentido, es recomendable que en los informes derivados de evaluaciones en las que utilizan el ACE se especifiquen las limitaciones asociadas a la medición realizada con los indicadores elegidos, así como los efectos indirectos, o sesgos, a los que su utilización pueda conducir. A la vista de la complejidad apuntada, para desarrollar buenas evaluaciones con el ACE es conveniente que las metodologías, y en concreto la definición de indicadores aplicables sectorialmente, sean debatidas y publicadas por instituciones

especializadas, tanto de la Administración pública (parece lógico que en España esta labor recaiga en la Intervención General de la Administración del Estado, IGAE, o en la futura institución de evaluación que se cree), como por instituciones profesionales y académicas. En ambos casos, las instituciones que lleven a cabo esta función deben tener en cuenta las experiencias de aquellos países considerados como referentes en la materia.

Finalmente, como ventaja añadida del ACE en relación con su utilización en el ámbito de la presupuestación por objetivos, debemos destacar su papel para conectar la evaluación y la presupuestación, en línea con lo expuesto en el artículo de Baños-Rovira (2023) publicado en el número 103 de *Ekonomiaz*. Para que el presupuesto por objetivos sea una herramienta útil para los gestores a la hora de promover la eficiencia económica, los créditos de gasto público deben clasificarse de forma que se permita conocer, sin ambigüedad, qué volumen de recursos se dedican a cada fin perseguido por los programas y proyectos financiados. Los indicadores empleados en el ACE, si están bien definidos, pueden vincularse claramente a los objetivos a conseguir en las políticas públicas, lo que permitirá evaluar la efectividad alcanzada por las políticas reflejadas en los presupuestos (Albi y Onrubia, 2018).

3.3. La medición de la eficiencia en las unidades gestoras

La medición de la eficiencia con la que operan unidades de decisión y producción públicas encargadas de producir los servicios públicos es una herramienta imprescindible para facilitar y mejorar su gestión. Además, su medida, en ámbitos de comparación homogéneos (p.e. servicios de cirugía traumatológica o cardiovascular de los hospitales de la red pública o de unidades de gestión de prestaciones de un determinado programa público) permite realizar clasificaciones de esos centros gestores del servicio en términos de su eficiencia o ineficiencia relativas. El análisis a partir de esas comparaciones de las causas por las que el desempeño o las prácticas seguidas hacen mejores o peores a unas unidades respecto de otras es fundamental para detectar ganancias potenciales de eficiencia y, en consecuencia, proceder a adoptar cambios organizativos, gerenciales o de otra índole, que permitan mejorar la gestión pública.

De acuerdo con la teoría económica, la eficiencia productiva de las unidades gestoras se puede contemplar desde dos vertientes: la técnica y la asignativa. La eficiencia técnica hace referencia a si se están utilizando de la mejor forma posible los factores productivos (por ejemplo, empleados públicos de cada categoría profesional,

inmuebles, equipos informáticos, elementos de transporte, etc.) para obtener el máximo nivel de actividad o desempeño posible (orientación de la eficiencia al *output*). Alternativamente, fijado un determinado nivel de actividad o desempeño, se puede analizar si se utilizan los niveles mínimos de esos factores (orientación de la eficiencia al *input*). De este modo, será eficiente técnicamente la unidad gestora que pueda producir la misma cantidad de bienes o servicios, con la calidad alcanzada, si reduce los factores productivos empleados. La elección de la orientación de la medición de la eficiencia productiva –al *input* o al *output*– debe depender, en buena medida, de la naturaleza de la actividad que se va a analizar. Por ejemplo, si se trata de problemas de calidad o de un servicio público sobre el que hay evidencias de problemas de cobertura, será preferible la orientación *output* del análisis, mientras que si se trata de analizar las posibilidades de reducir costes sin deterioro de las condiciones de prestación del servicio, la orientación *input* será la adecuada.

Existe eficiencia asignativa si, considerando unidades gestoras técnicamente eficientes, se utiliza la combinación de *inputs* que ofrece el coste económico mínimo, o lo que es lo mismo, si dado un volumen de gasto, la unidad gestora maximiza el valor de los bienes y servicios que provee. No se trata de ver si se utiliza bien o no una tecnología de producción (algo asociado con la eficiencia técnica), sino que la tecnología de producción del servicio sea la más eficiente, esto es, si la unidad está obteniendo el servicio público analizado sin emplear una cantidad excesiva de algún factor y demasiada poca de otro, teniendo en cuenta los precios respectivos de esos factores y sus productividades marginales. Por último, el concepto económico de eficiencia o eficiencia productiva global es el resultado de la combinación de la eficiencia técnica y de la asignativa. Intuitivamente, la eficiencia global representa el aprovechamiento pleno de la productividad de cada uno de los factores productivos (eficiencia técnica), así como la idoneidad con la que estos deben combinarse en el proceso productivo de acuerdo con sus productividades marginales y con sus respectivos precios unitarios (eficiencia asignativa).

En el terreno aplicado, la realización de análisis de eficiencia productiva se concentra, de forma muy mayoritaria, en el ámbito de la eficiencia técnica. Esto se explica porque para desarrollar análisis de eficiencia asignativa es necesario disponer de precios unitarios de los *inputs*, algo que no es tan sencillo en muchas actividades públicas. No obstante, hay que precisar que en los últimos años se está extendiendo un enfoque, en cierta medida alternativo, que trata de solventar las carencias de las mediciones exclusivas de eficiencia técnica, desconectadas, al menos de forma directa, con los potenciales ahorros presupuestarios. Se trata de una literatura de

carácter fundamentalmente empírico que se centra en analizar directamente la eficiencia del gasto público, considerando este como un *input*, utilizando como medidas de *output* diferentes indicadores de desempeño de las actividades públicas a evaluar.¹⁴ En su aplicación original, este enfoque de la eficiencia del gasto público propuesto en Afonso *et al.* (2005) contempla comparaciones de las principales políticas públicas entre países. En la actualidad se realizan aplicaciones para el gasto regional que comparan la eficiencia en el gasto entre gobiernos regionales.¹⁵

La eficiencia técnica se puede medir mediante diferentes metodologías. Básicamente, estas se dividen en dos grupos atendiendo al tipo de tratamiento cuantitativo que se dé a los datos. Así, por un lado, tenemos los métodos no paramétricos, cuyas técnicas utilizan fundamentalmente la programación lineal; y por otro, los métodos paramétricos, que emplean estimaciones econométricas. Los métodos no paramétricos se utilizan con especial profusión desde los años ochenta del siglo XX, especialmente en el campo de la educación y de la sanidad pública, aunque podemos encontrar aplicaciones para prácticamente todos los servicios públicos. Dentro de estos, la técnica más utilizada es el conocido como Análisis Envolvente de Datos (DEA, por sus iniciales en inglés) y su variante el FDH (Free Disposal Hull).¹⁶ Estas técnicas no paramétricas no requieren establecer *a priori* la forma específica de la función de producción que relaciona *inputs* y *outputs*, lo que supone una ventaja para el análisis aplicado ya que evita la explicitación de la tecnología productiva empleada por las unidades gestoras comparadas. Otra ventaja son los supuestos poco restrictivos que requiere la estimación de las fronteras de producción no paramétricas, las cuales se obtienen a partir de las unidades gestoras que se comportan mejor en cuanto al rendimiento de sus *inputs* empleados o, si consideramos una orientación al *output*, de las unidades que se comportan mejor de acuerdo con los *inputs* utilizados. Como principal limitación de estos métodos no paramétricos hay que apuntar la imposibilidad de predecir, a partir de los resultados, las cantidades máximas de *output* derivadas de los cambios a adoptar en las

¹⁴ En Afonso *et al.* (2010) se extiende este tipo de análisis de eficiencia del gasto público a políticas con objetivos redistributivos. Recientemente, en Afonso *et al.* (2023) se ha recopilado un conjunto de trabajos que revisan el estado actual de esta literatura, incluyendo experiencias internacionales diversas, que muestran las potencialidades de la aplicación práctica de este enfoque en diferentes ámbitos de las actuaciones del sector público.

¹⁵ García-Gómez *et al.* (2019) evalúan la eficiencia del gasto en educación y sanidad, entre los años 2005 y 2015, en las comunidades autónomas españolas.

¹⁶ La técnica DEA define una frontera como combinación lineal de las unidades eficientes, mientras que el FDH lo hace exclusivamente con las observaciones de las unidades eficientes. De este modo, el FDH permite una interpretación comparativa con las unidades ineficientes más sencilla e intuitiva, al no hacerse con unidades "artificiales" representadas en los puntos de la combinación lineal de las unidades eficientes.

cantidades empleadas de *inputs* o, alternativamente, las necesidades de *inputs* requeridas para alcanzar niveles deseados de *output*.

Una alternativa son los métodos paramétricos basados en la estimación de modelos econométricos de naturaleza estocástica. A diferencia de los primeros, estos modelos necesitan una especificación funcional concreta, bien sea la correspondiente a la función de producción, bien la referida a la función de costes. En cualquiera de las dos alternativas se trata de ecuaciones a estimar econométricamente. Estos métodos miden, básicamente, la ineficiencia técnica a partir de los residuos de la estimación econométrica, de ahí que el resultado de su aplicación sea conocido como estimación de "fronteras estocásticas". En Santín (2015) puede consultarse una presentación bastante intuitiva de estos modelos, diferenciando entre los modelos aplicables a paneles de datos (Battese y Coelli, 1992), que tienen en cuenta variables de entorno explicativas de la ineficiencia (Battese y Coelli, 1995) y los modelos con múltiples *inputs* y *outputs* (Coelli y Perelman, 1999).

Una limitación de estos métodos es que debe tenerse en cuenta que en el mundo real las fronteras de producción pueden cambiar a lo largo del tiempo. Los índices de Malquist (Caves *et al.*, 1982) son una interesante vía para poder valorar esos cambios y analizar la evolución de la productividad de los factores. Además, estos índices permiten descomponer los efectos sobre la productividad de los factores debido al progreso tecnológico y a los aumentos de la eficiencia.

En cuanto a la utilidad de estas técnicas, hay que tener en cuenta que la clasificación de las unidades gestoras de acuerdo con su nivel de eficiencia requiere, además, un análisis de las causas que explican que unas sean más eficientes que otras, siempre con la referencia de la frontera establecida. De hecho, las causas de la ineficiencia productiva pueden ser muy variadas: a) el marco legal que afecta a cada actividad pública puede condicionar los resultados; b) los objetivos a alcanzar pueden no estar suficientemente explicitados; c) los esquemas de incentivos que afectan a la gestión pueden ser limitados; d) los diseños organizativos en los que se gestionan las actividades; e) la productividad de los empleados públicos puede estar condicionada por las limitaciones del sistema de selección y retribución aplicado, lo que estaría generando un factor más de ineficiencia, al que Leibenstein (1966) denominó "ineficiencia X".

En todo caso, el estudio y análisis de las causas que están detrás de la ineficiencia en la gestión de las diferentes áreas de provisión de servicios públicos es, en nuestra

opinión, una herramienta imprescindible para impulsar una mayor productividad de las actividades públicas. Desde la perspectiva de la institucionalización de la evaluación, esta vertiente del análisis de eficiencia productiva constituye un elemento fundamental, a través de la información generada, para que puedan alcanzarse satisfactoriamente los fines esperados de la evaluación. Debemos advertir aquí que esto requiere cumplir con una serie de condicionantes, como son: a) la selección de las metodologías adecuadas; b) la formación especializada de los empleados públicos encargados de aplicar estas técnicas; c) el desarrollo y seguimiento temporal y periódico de este tipo de análisis. La creación de unidades de control interno en los centros gestores, generadoras de información de calidad, es, igualmente, un requisito fundamental (Onrubia, 2005).

3.4. La evaluación de impacto

Como señala Moral-Arce (2023) en su artículo del número 103 de *Ekonomiaz* en el que presenta las potencialidades que la evaluación de impacto tiene en relación con la calidad de las finanzas públicas, esta constituye un elemento esencial: "Si se desean políticas públicas que logren solucionar los problemas de la sociedad es necesario, por un lado, generar evidencia para una mejor toma de decisiones, y por otra parte, que estos análisis estén orientados a la obtención de resultados, y la evaluación de impacto cubre ambos objetivos".

La finalidad de la evaluación de impacto es tratar de comprobar si una política pública, bien sea una medida de intervención concreta, un programa o un proyecto, consigue los resultados esperados que justificaron su puesta en marcha. De acuerdo con su enfoque metodológico, la evaluación de impacto estudia la efectividad que intervenciones públicas concretas tienen para sus potenciales beneficiarios, midiendo los cambios recogidos en uno o en varios indicadores tras su aplicación, tratando de aislar los efectos que sobre esos indicadores han podido generar otras variables o circunstancias no directamente relacionadas con la actividad evaluada.

Para realizar una evaluación de impacto existen diversas técnicas cuantitativas, cuyo rasgo común es determinar relaciones de causalidad entre la actuación considerada y sus efectos sobre las personas, las empresas o los mercados. Se trata, por tanto, de técnicas dirigidas a proporcionar mediciones del impacto causal que la aplicación de un proyecto, un programa o un cambio regulatorio tiene en la sociedad, es decir, en general, el impacto en el bienestar social de una determinada política pública.

Como es sabido, la determinación de la causalidad en las relaciones cuantitativas entre variables hace necesario diferenciar dichas relaciones de las simples correlaciones entre variables. Para establecer efectos causales que informen con precisión y sin ambigüedad de que un cambio en una variable determinada causa un cambio en otra variable, lo ideal es disponer de datos generados en experimentos naturales totalmente aleatorios.¹⁷ Este tipo de análisis experimental es habitual en ámbitos como la medicina, donde la efectividad de los nuevos medicamentos o procedimientos médicos es evaluada mediante la comparación del estado de salud de individuos incluidos en un grupo a los que se dispensa un nuevo tratamiento, denominado "grupo de tratamiento", y el estado de salud de los individuos no tratados, que forman el denominado "grupo de control" o "grupo contrafactual", cuyos miembros deben tener las mismas características que los pertenecientes al grupo que recibe el nuevo tratamiento. La aleatoriedad exige que los candidatos sean asignados al azar para formar parte de ambos grupos.

No obstante, hay que advertir que estos experimentos no son fáciles de realizar en la mayor parte de los escenarios de aplicación de las políticas y programas públicos, donde normalmente se carece de "laboratorios" experimentales. Frente a la imposibilidad de contar con experimentos naturales, surge la necesidad de diseñar experimentos o cuasi-experimentos sociales en los que conforme a la información estadística disponible se determina los individuos (u hogares o empresas) que forman parte del grupo "tratado" por una intervención pública y los no afectados por dicho programa, que engrosan el grupo de control. Para ambos colectivos se obtienen indicadores de los resultados a lo largo del tiempo, con la condición de que sean comparables entre sí. Para que exista aleatoriedad a la hora de confeccionar los dos grupos, la selección de los individuos no debería verse afectada por ninguna otra causa que no sea la de los efectos del programa evaluado, algo que en la práctica resulta bastante difícil de conseguir. Ante esta limitación, aparecen los conocidos sesgos de selección muestral que representan las diferencias, no siempre observables, entre los miembros de los grupos de tratamiento y control. Esto provoca que esas diferencias afecten espuriamente tanto a la participación de los individuos en los programas públicos evaluados como a los propios resultados obtenidos en la evaluación, aunque en la realidad no se deban a su aplicación.

Las técnicas de inferencia causal que se aplican en las evaluaciones de impacto pueden agruparse en dos categorías, basadas en el tipo de diseño experimental que

¹⁷ Las características deseables que deberían reunir los experimentos y cuasi-experimentos naturales en la evaluación de impacto pueden verse en DiNardo (2008).

se vaya adoptar, con el propósito de mitigar la falta de aleatoriedad apuntada. Así, según la forma en la que los participantes son asignados a los grupos de tratamiento y de control, se distingue entre diseños experimentales aleatorios, generalmente difíciles de conseguir, y los denominados diseños cuasi experimentales, de carácter no estrictamente aleatorios. En ambos casos, estos experimentos sociales son, por regla general, complejos de desarrollar y su coste de implementación depende, en buena medida, de la disponibilidad de datos estadísticos de calidad, algo que va a depender de si en el propio diseño de la política que se desea evaluar se ha contemplado o no un procedimiento adecuado de recogida de información estadística. Como se expone en Moral-Arce (2023), cada técnica de evaluación de impacto tiene sus propias exigencias de información estadística.

Las técnicas actualmente utilizadas en la evaluación de impacto poseen una clara ventaja sobre las técnicas de evaluación tradicionales en lo que concierne a su capacidad para inferir la causalidad en las relaciones analizadas. Esto es debido a que las técnicas tradicionales se limitan, por regla general, a hacer comparaciones del tipo “antes-después” con respecto a la intervención que se quiere evaluar, pero aplicando el análisis cuantitativo solo con muestras integradas por individuos que han recibido el tratamiento de la política, o en comparaciones del tipo “con-sin” tratamiento, pero solo con datos posteriores a la aplicación de la política (Santín, 2015). De hecho, hasta hace unas dos décadas, una gran cantidad de los análisis realizados para evaluar el impacto de políticas públicas medían los pretendidos efectos causales mediante series temporales o de regresiones de sección transversal, sin un diseño formal del experimento social que garantizase una mínima aleatoriedad.

En todo caso, debe advertirse que, a pesar de los avances producidos en el campo del diseño de experimentos sociales, la práctica de evaluaciones de impacto puede verse afectada por sesgos ante la presencia de variables inobservables (como la motivación intrínseca o los incentivos) que afecten simultáneamente a la participación de los individuos en un programa y a los resultados producidos. También pueden existir sesgos como consecuencia de otras variables no observadas afectadas por cambios en el periodo estudiado y que incidan tanto en los individuos asignados al grupo de tratamiento como en los del grupo de control. Frente a ello, los sesgos de selección muestral pueden ignorarse si, por ejemplo, en un análisis temporal se detecta una fuerte alteración, no prevista, en las tendencias de las series, la cual puede coincidir con un cambio importante en una variable. En ese caso, podríamos interpretar que esa variable ha originado un efecto causal en el resultado. Mediante el análisis de regresión transversal (referidos a un momento temporal concreto), la

elección de variables de “control” que permitan dividir la muestra en dos grupos que, en el resto de características sean iguales, puede ser una alternativa para determinar de forma razonable la existencia de una relación causa-efecto como consecuencia de la participación en el programa evaluado, lo que permitiría ignorar los sesgos de selección comentados.

En relación con la utilización de la cuasi experimentación en las evaluaciones dirigidas a evaluar el impacto de las políticas públicas, las técnicas de inferencia causal utilizadas están orientadas a minimizar los mencionados sesgos de selección.¹⁸ Entre ellas, podemos destacar la denominada “regresión en discontinuidad”, que se suele aplicar cuando los beneficiarios potenciales de un programa público se determinan según un punto de corte en alguna variable que sea factible de observar (renta familiar, edad, número de dependientes, etc.). Ese corte de la continuidad permite formar los dos grupos del análisis (por ejemplo, disponer de una renta familiar inferior o superior a una cuantía concreta formará, respectivamente, los grupos de “tratados” y no “tratados”), suponiendo que en todas las demás características los integrantes de grupos son iguales o similares. En el supuesto de que las características de los individuos cercanos al punto de corte sean parecidas, puede realizarse una comparación de los resultados medios de los individuos de ambos grupos que se sitúan en el entorno del punto de corte.

Asimismo, esta clase de técnicas son utilizadas cuando en el diseño de un experimento para evaluar una política existen indicios de que los grupos de tratamiento y de control no se han formado con suficiente aleatoriedad. Por ejemplo, en el caso de evaluar el impacto que un programa de garantía de rentas mínimas sobre la oferta de trabajo puede tener en una región (grupo de tratamiento) y no en otra (grupo de control), admitiendo que la situación económica de ambas es similar. En ese caso, para contrastar si las muestras están o no sesgadas es recomendable, si los datos lo permiten, comparar los dos grupos en dos periodos de tiempo distintos, antes y después de la aplicación del programa que se evalúa. No obstante, hay que advertir que también en la segunda comparación puede existir sesgo de selección, en la medida que los grupos nos son formados de manera estrictamente aleatoria.

El método de “diferencias en diferencias” es una opción a considerar cuando hay sesgos de selección de esta naturaleza. Este método calcula la diferencia de los

¹⁸ Además del artículo de Moral-Arce (2023) en el número 103 de *Ekonomiaz*, los manuales de Khandker *et al.* (2010) y Gertler *et al.* (2011), publicados por el *World Bank*, contienen completas guías prácticas para la aplicación de las principales técnicas de evaluación de impacto. El libro coordinado por García-Pérez (2011) es una buena guía de aplicación de estas técnicas a diferentes tipos de políticas, incluido el diseño de cuasi experimentos y alternativas para tratar de mitigar los sesgos de selección.

valores medios de la variable que recoge el efecto de la aplicación del programa, antes y después de su aplicación, tanto en el grupo de tratamiento como en el de control, procediendo en una segunda fase a calcular la diferencia de esas dos primeras diferencias. La segunda diferencia estaría así midiendo el impacto del programa evaluado, en la medida que con ella se observa el cambio que se ha producido, antes y después de la entrada en vigor del programa, en cada uno de los dos grupos (primera diferencia), aislándose en el cálculo de la segunda diferencia aquellos factores omitidos que podrían estar sesgando el análisis transversal entre ambos grupos debido a la selección no aleatoria de las observaciones. Como señala Santín (2015), esta metodología implica aceptar que la tendencia del sesgo inicial (antes de aplicar la política) respecto del indicador de resultados utilizado en los dos grupos, es constante en el tiempo y desaparece en las diferencias calculadas. No obstante, el método de "diferencias en diferencias" tampoco asegura que los resultados que proporcionan carezcan de sesgo, pues el supuesto anterior no siempre tiene por qué cumplirse.

Otras técnicas utilizadas en la evaluación de impacto son el Propensity Score Matching (PSM) y la incorporación de variables instrumentales (IV) en los análisis econométricos. El método PSM consiste en construir estadísticamente los grupos de comparación a partir de un modelo de probabilidad de participación en el programa (o de ser tratado por la medida evaluada), utilizando para ello características observables de los individuos. Por ello, su utilización es recomendable en la evaluación de programas públicos en los que existe elevada verosimilitud de que la participación en los mismos depende exclusivamente de las características observables de los agentes. A partir de las probabilidades de participación y no participación (*propensity score*) los individuos se emparejan (*matching*) creando una muestra estadística en la que los efectos medios generados por el programa se calculan como una media de las diferencias entre los dos grupos originada en la variable seleccionada para medir el impacto individual de la medida (Harris y Horst, 2019). Una complicación para el uso de este método es que requiere definir un tipo concreto de *matching* estadístico para emparejar a los individuos tratados y no tratados por la medida.¹⁹ En los últimos años, el método PSM puede decirse que ha perdido presencia en la evaluación de impacto. Esto puede deberse a la exigencia que imponen dos condiciones, apuntadas en Khandker *et al.* (2010): a) que los factores no observables no afecten a la participación (existencia de independencia condicional); b) que las muestras de participantes y no participantes permitan

¹⁹ Sobre los métodos de fusión estadística de bases de datos para el PSM, véase el trabajo de Caliendo y Kopeinig (2008).

disponer de un conjunto suficientemente amplio de observaciones para obtener estimaciones consistentes de las probabilidades de participación.

En relación con los métodos que incorporan “variables instrumentales” (IV), la literatura señala que son recomendables cuando las evaluaciones se enfrentan a problemas de endogeneidad vinculados a la participación de los individuos en los programas de acuerdo con las condiciones de elegibilidad establecidas, dando lugar a resultados sesgados. El método de variables instrumentales (Imbens y Angrist, 1994; y Angrist *et al.*, 1996) afronta, de forma directa, la selección de variables no observables a través de la introducción en las estimaciones de una variable instrumental, que afecta exógenamente al proceso de selección pero sin estar directamente relacionada con el resultado o con las variables no observadas que determinan este. Como destacan Blundell y Costa-Dias (2009), con la incorporación de variables instrumentales se introduce una fuente de variación exógena que potencia la aleatorización en la asignación de los participantes a los grupos de tratamiento y control. La complejidad principal del método IV estriba en la selección del instrumento adecuado, donde además de su plena exogeneidad, la elección debe estar fundamentada por criterios económicos claros y altamente correlacionada con los elementos determinantes de la participación en el programa evaluado. La utilización del método es especialmente adecuada en los análisis econométricos con datos de panel, siendo la regresión en dos etapas el procedimiento habitualmente más utilizado.

En resumen, la práctica habitual de las evaluaciones de impacto de las políticas públicas, como se destaca en el artículo de Moral-Arce (2023), requiere el diseño de cuasi experimentos sociales en los términos que acabamos de exponer. No obstante, debe advertirse que, como cualquier otro método, su aplicación no está exenta de limitaciones. En este sentido, para una adecuada realización, cada evaluación requiere un estudio previo en el que, de acuerdo con su problemática y en línea con lo apuntado, se valoren los pros y los contras que apareja la elección de una técnica de análisis concreta, así como la complejidad del experimento social a realizar. Por ello, la realización sistemática de este tipo de evaluaciones de impacto requiere contar con especialistas en las técnicas a emplear, así como seguir los métodos y procedimientos asentados en la literatura, y que son considerados como buenas prácticas.

4. El marco institucional de la evaluación de políticas públicas

Una vez presentadas las principales herramientas disponibles para llevar a cabo la evaluación económica de proyectos y programas públicos, en esta sección nos centramos en la manera en la que esta, a la luz de las mejores experiencias comparadas, debería implantarse en los diferentes procesos que articulan la gestión pública, considerada en un sentido amplio, desde la planificación de las actividades a desarrollar por el sector público hasta la valoración de los resultados de su realización. Para ello, es necesario considerar el contenido y alcance de cada tipo de evaluación, las técnicas disponibles para cada uno, así como la identificación de los responsables de su realización y de los destinatarios de sus resultados, incluida la fijación transparente de las consecuencias asociadas a estos. Debe advertirse, por obvio que pueda parecer, que la simple aplicación de las técnicas no garantiza que lo hayan sido de forma adecuada, ni que aun siéndolo, los resultados de las evaluaciones prácticas vayan a rendir los efectos de mejora de la calidad de las finanzas públicas que se espera. En este sentido, asegurar este aprovechamiento es el papel fundamental que debe desempeñar el marco institucional de la evaluación económica.

Respecto al marco institucional, una primera cuestión que debemos tener en cuenta es que no deben confundirse las técnicas de análisis y evaluación con el diseño de los procesos sistemáticos, más o menos reglados, en los que se aplican estos instrumentos. Así, los procesos de evaluación y control, ya operen *ex ante*, con un carácter prospectivo respecto de las decisiones de gasto público, o *ex post*, con un carácter supervisor, deben incorporar las técnicas más adecuadas para los fines perseguidos, teniendo en cuenta su contenido y métodos de aplicación.

Cuando nos fijamos en estos procesos, de acuerdo con sus diferentes fines, la literatura –tanto académica como profesional e institucional– se refiere a términos como las auditorías de gestión (por su contenido, quizás el proceso de evaluación con mayor amplitud de fines), las auditorías de eficiencia o de eficacia (*performance audit*) o a las denominadas auditorías *value for money*, de realización generalizada en los países anglosajones y del norte de Europa.

El diseño de los diferentes procedimientos para evaluar la actuación pública desde una perspectiva económica supone, a su vez, elegir las técnicas más idóneas de acuerdo con los fines perseguidos. Así, si lo que se persigue es seleccionar las

alternativas de gasto más eficientes o con mayor rentabilidad social para cubrir una necesidad social, parece lógico contemplar la instauración sistemática de procesos de evaluación de proyectos, con incorporación de técnicas como el ACB o el ACE. Si de lo que se trata es de realizar un seguimiento del cumplimiento de los objetivos vinculados a un determinado programa público, la realización sistemática de auditorías de cumplimiento de objetivos (*performance audit*) es una medida apropiada de diseño institucional. Para conseguir una rendición de cuentas con una efectiva exigencia de responsabilidades de gestión, la implementación institucional de auditorías tipo *value for money*, con análisis comparativos de eficiencia productiva y de evaluación de impacto, se convierte en una exigencia fundamental.

4.1. Aspectos teóricos del diseño institucional

Como señalan Debrun *et al.* (2009), en las actividades de control, supervisión o evaluación de las finanzas públicas, para garantizar su efectividad y, en definitiva, mejorar su calidad, la independencia en el ejercicio tanto de las tareas técnicas como de las directivas resulta un elemento esencial. La literatura académica reconoce desde hace mucho tiempo que los políticos actúan de forma bastante racional en relación con sus decisiones presupuestarias, movidos por incentivos específicos asociados con su reelección, las presiones de los grupos de interés, sus compromisos electorales u otros acuerdos asumidos en el ejercicio de sus funciones políticas. En consecuencia, resulta bastante obvio que la ausencia de independencia respecto del poder político en las instituciones que tienen encomendado el desarrollo de estas actividades condiciona en el origen la validez de las mismas, en cuanto a que sus resultados quedan afectados por la falta de credibilidad.

Una cuestión muy relevante en la institucionalización de la evaluación del gasto público tiene que ver con la terminología y el alcance y contenido de los procedimientos que se suelen utilizar en los distintos países para su realización. Tanto en la legislación que establece los diversos tipos de controles como en los manuales de procedimientos seguidos por las instituciones responsables del control económico-financiero, es frecuente encontrar referencias a diferentes tipos de auditoría: de gestión –quizás el concepto más empleado–, operativa, de resultados, integral, de eficiencia y eficacia, *value for money*, incluso auditoría organizativa o auditoría de administración. Para Parker (1986) se trata, en gran medida, de una mera diferenciación nominal, pues en el fondo realiza análisis valorativos de la gestión realizada desde un enfoque eminentemente económico, diferenciado del estrictamente vinculado a la regularidad financiera y legal. Por ello, Parker (1986) considera que el concepto más adecuado es el de auditoría de gestión, pues recogería

el análisis integral de la gestión realizada por una agencia o entidad gestora, incluyendo la evaluación de su funcionamiento organizativo y de su rendimiento en términos de desempeño en cada una de sus áreas operativas y de actividad, bajo una aproximación económica, con la eficiencia y la eficacia como criterios rectores.

Sobre esta concepción, en la década de los ochenta y en el ámbito del sector público se acuñó en Estados Unidos el término *performance audit*, con un cierto paralelismo al *operational auditing*, ya utilizado en ese país, pero en el ámbito de las entidades privadas, y que englobaba tanto aspectos financieros como económicos y organizativos.

Sin embargo, a pesar de la indudable relevancia de los aspectos conceptuales para el propio desarrollo técnico de estas auditorías, estamos de acuerdo con Kells y Hodge (2009) cuando advierten de que una completa definición del concepto de *performance audit* (y por extensión, de la auditoría de gestión) no debería depender exclusivamente de la metodología de desarrollo o de la forma de analizar los criterios de economía, eficiencia y eficacia (las conocidas como "tres Es"), sino de aspectos institucionales como la independencia del poder gubernamental de la entidad encargada de su realización, su adscripción parlamentaria, la selección y el nombramiento de los profesionales que deben realizar las auditorías, el alcance de sus evaluaciones en cuanto a la rendición de cuentas y la exigencia de responsabilidades, así como la transparencia en la divulgación de los resultados alcanzados y los protocolos seguidos.

En esta orientación estamos de acuerdo con Johnsen *et al.* (2001), para quienes la auditoría de gestión es un instrumento recomendable para mejorar el funcionamiento democrático de las instituciones políticas, pues no se trata solo de un instrumento para potenciar el uso eficiente y eficaz de los recursos económicos, sino también de un elemento fundamental para hacer efectiva la rendición de cuentas. De hecho, esta es la filosofía subyacente al concepto *value for money auditing*, asentada desde hace décadas en los países nórdicos y anglosajones, donde se vinculan con claridad los resultados de estas evaluaciones de gestión con la exigencia de responsabilidades.

Por lo que respecta a la capacitación y especialización de los profesionales que han de encargarse de realizar estas auditorías y evaluaciones, existe un amplio consenso en la literatura sobre la necesidad de contar con equipos multidisciplinares en los que los economistas desempeñen un papel preponderante, pero en absoluto exclusivo (Glynn, 1985). La gran diversidad y complejidad de las actividades realizadas por el

sector público obliga a incluir en los equipos a especialistas de las distintas áreas, pues realmente no puede ser concebida como una práctica auditora en sentido estricto, sino un instrumento de evaluación tendente a promover la adecuación del uso de los recursos públicos a criterios económicos (Barzelay, 1997).

Asimismo, la realización de evaluaciones de impacto de las medidas regulatorias se ha convertido en una herramienta muy valiosa en los programas de regulación de los principales países de la Unión Europea. No obstante, existen diferencias notables entre países en cuanto a la incorporación regular de este tipo de evaluaciones a la toma de decisiones gubernamentales, dependiendo del grado de éxito en su utilización, de la capacidad administrativa de los gobiernos, así como de los recursos económicos empleados en las actividades de evaluación y, sobre todo, de la voluntad política respecto de la publicación de los resultados de estas evaluaciones (De Francesco *et al.*, 2012).

Blume y Voigt (2011) estudian la relevancia del diseño organizativo de las instituciones superiores de control en un trabajo empírico para más de cuarenta países y concluyen que aquellas instituciones que se organizan como "agencias auditoras" presentan una mayor efectividad en la detección de malas prácticas con incidencia en la corrupción. Como factores cruciales para el funcionamiento con éxito de estas instituciones se señalan: a) la claridad y determinación del mandato realizado (diferenciando con claridad el tipo de auditorías: financieras, de cumplimiento o *value for money*); b) el compromiso de publicitación del informe y su completitud; c) la explicitación de las consecuencias derivadas de un informe de evaluación negativo; d) la independencia efectiva y real de la institución de supervisión, incluida la capacidad para recabar información de los centros supervisados; e) la independencia real de los técnicos auditores y evaluadores, que no deberían mantener lazos profesionales estrechos presentes ni futuros con la administración pública; f) la autonomía para fijar las retribuciones y carreras profesionales de su plantilla; g) la disponibilidad suficiente de financiación presupuestaria; h) el marco institucional que afecta a la exigencia de responsabilidades, incluido el régimen de transparencia; i) la libertad de prensa y la fortaleza de la sociedad civil. El reconocimiento constitucional del mandato de supervisión constituye un reforzamiento añadido de estas instituciones, si bien este requisito, de forma aislada, no asegura el funcionamiento correcto de las mismas (Frey, 1994).

En relación con las dificultades que presentan muchos países para conseguir un funcionamiento adecuado de estas instituciones de control y evaluación, Haiarinejad *et al.* (2012) señalan como un problema bastante común la falta de compromiso firme de los gobiernos con la implantación efectiva de los sistemas de evaluación. Las deficiencias de los sistemas de control interno de los centros gestores y las limitaciones informativas de la gestión presupuestaria, con escasa implantación de una verdadera presupuestación por objetivos, condicionan fuertemente la realización de evaluaciones rigurosas de los programas de gasto. En muchos casos, más allá de la aprobación del marco legal y de la constitución formal de estas instituciones, no existen normas suficientes sobre los programas de auditoría de resultados.

Además de los factores enumerados, la voluntad política (*political will*) constituye un factor clave en el correcto funcionamiento de la evaluación del gasto público. Como se desprende de los requisitos expuestos, el cumplimiento de muchos de ellos depende en gran medida de decisiones a adoptar, bien por los propios centros supervisados, bien por los departamentos gubernamentales de los que dependen; en definitiva, del poder ejecutivo. Esta situación hace que la independencia del gobierno, si bien constituye para la teoría, como hemos visto, un requisito imprescindible, sea una condición necesaria pero no suficiente. Aquí es donde la *political will* emerge como un valor fundamental. Como destaca Heald (2012), el papel de la transparencia es, sin duda, capital para desarrollar esta voluntad política, si bien esta no es fácil de conseguir en relación con el gasto público dada la complejidad de los objetivos que persiguen los programas de gasto y su fuerte naturaleza política.

4.2. Mejores prácticas internacionales en evaluación pública²⁰

4.2.1. Alemania²¹

En Alemania, la evaluación de los programas de gasto público la realiza principalmente el Tribunal de Cuentas Federal (Bundesrechnungshof, BR), que cuenta con una estructura organizativa con más de 1.400 empleados públicos. El BR es una institución adscrita al Parlamento, independiente del Gobierno. El BR tiene encomendado constitucionalmente examinar la gestión financiera federal, incluyendo en su mandato de auditoría también a la seguridad social y a las empresas privadas en las que participa el sector público alemán, así como a las privatizadas.

²⁰ Este apartado recoge, de forma sintética, el contenido del trabajo de investigación publicado en Albi y Onrubia (2016) y desarrollado posteriormente en Albi y Onrubia (2018). También se incorpora información aportada en Novales (2023) para Francia y Países Bajos.

²¹ https://www.bundesrechnungshof.de/EN/0_home/home_node.html

El BR remite anualmente un informe a las dos Cámaras del Parlamento alemán y al Gobierno federal que es utilizado como base para la aprobación por el Parlamento de la gestión del Ejecutivo. Este informe se presenta en público, con asistencia de los medios de comunicación.

Las actividades de supervisión que realiza el BR son de amplio contenido y abarcan desde las tradicionales auditorías financieras de cuentas y de regularidad a las auditorías de gestión, fundamentalmente auditorías de desempeño, las cuales han cobrado un papel preponderante en los últimos años. Además, el BR ha potenciado de forma notable sus labores de asesoramiento a las entidades auditadas, en base a los resultados de sus trabajos de auditoría integral, destacando sus actividades de consultoría con recomendaciones para la mejora de la calidad de los servicios públicos, así como para analizar los potenciales ahorros de gasto en las unidades gestoras o la posibilidad de aumentar los ingresos para financiar las actividades. Asimismo, el BR también realiza informes de proyectos de ley y de los principales proyectos de gasto público, ya sea de forma puntual o en la elaboración anual de los presupuestos.

En los últimos años, el BR se encuentra en un proceso de cambio de orientación de su actividad, con el objetivo de adecuar más sus actividades hacia la realización de evaluaciones de gestión, compensando así una tradicional orientación hacia el control contable y financiero. Las auditorías de desempeño (*performance audits*) suelen centrarse en evaluar si la actividad de los centros gestores de gasto tiene lugar bajo criterios de eficiencia y eficacia, valorando si la provisión de servicios públicos presenta una buena relación calidad-precio. Sus informes también prestan una especial atención a los recursos humanos empleados en las actividades públicas, así como a la capacidad del sector público para cubrir los objetivos asignados a sus actuaciones. De hecho, la evaluación de programas y el análisis del cumplimiento de sus objetivos se han extendido con bastante efectividad a los programas federales más importantes, así como a los proyectos del Gobierno con mayor trascendencia económica y social.

4.2.2. Australia²²

La Australian National Audit Office (ANAO) es la institución encargada de realizar auditorías de las entidades del sector público australiano. La ANAO es una entidad adscrita al Parlamento y, por tanto, independiente del Gobierno. Su máximo

²² <https://www.anao.gov.au>

responsable, el Auditor General, es nombrado por el Primer Ministro por un mandato de diez años, siendo su puesto inamovible.

De acuerdo con la cultura de servicio a la sociedad, establecida como un objetivo en la refundación de la ANAO en 1997, se considera al Parlamento australiano como cliente principal de sus actividades. Estas tienen como cometido ofrecer evaluaciones independientes de las áreas seleccionadas de la administración pública, así como proporcionar informes financieros del sector público, análisis organizativos de la administración y facilitar la rendición de cuentas. También son considerados clientes relevantes el Gobierno y las entidades que integran el sector público, en la medida que los informes de las auditorías realizadas han de incluir recomendaciones de mejora de la gestión, desde principios de eficiencia y eficacia, así como de mejora de la calidad de los servicios públicos.

En este marco institucional, la ANAO asume el compromiso de realizar una evaluación objetiva y constructiva de las áreas donde se considera que pueden hacerse mejoras en la administración pública y en la prestación de servicios. Para llevar a cabo su misión, la ANAO debe desarrollar sus actividades bajo un principio de cooperación entre los responsables de las entidades gestoras y los comités de auditoría encargados de las supervisiones. La ANAO asume también tareas de consultoría y de elaboración de guías de “mejores prácticas” de gestión en las diferentes áreas de la administración pública.

En el ámbito del desarrollo de normas de auditoría y de formación de profesionales, la ANAO cuenta con una tradición prestigiosa tanto a nivel nacional como internacional. En este sentido, son frecuentes los intercambios de experiencias entre expertos nacionales e internacionales, así como entre organizaciones profesionales. En relación con su personal, la ANAO ofrece procesos de selección abiertos a profesionales de la auditoría y la evaluación, ofreciendo atractivas carreras profesionales dentro de su organización.

La ANAO realiza tres tipos de actuaciones: auditorías de gestión, auditorías de estados financieros y revisiones de calidad (*quality assurance reviews*). En cuanto a su funcionamiento, las actuaciones de la ANAO se rigen por los estándares profesionales de las normas de auditoría aplicados en Australia. Con ello, se trata de asegurar que los informes de auditoría de gestión, de auditoría de estados financieros y de revisión de la calidad están diseñados para ofrecer un nivel razonable de fiabilidad, objetividad y rigor. En todo caso, se reconoce expresamente que el grado

de fiabilidad de las actuaciones auditoras vendrá determinado, en cada caso, por factores como el objetivo concreto de la auditoría, las limitaciones inherentes del control interno existente en las entidades auditadas, así como la disponibilidad documental y los costes de realización de las tareas de supervisión.

Las áreas de actividad seleccionadas son incluidas en un programa plurianual de auditoría, de acuerdo con las prioridades establecidas por el Comité Conjunto de Cuentas Públicas y Auditoría del Parlamento, que aprueba el Auditor General. En esta selección son tenidas en cuenta, normalmente, las opiniones de las entidades y de otros agentes involucrados en las actividades objeto de supervisión. El diseño de este programa aspira a una cobertura lo más amplia posible de las diferentes áreas de actividad del sector público, discriminando en función del riesgo, incluido el volumen de recursos gestionado, y la relevancia social de las actividades. En aplicación del principio de transparencia asumido por la ANAO, todos los informes realizados son accesibles públicamente a través de su página web, donde figuran ordenados temporalmente por tipo de informe, por área temática y actividad.

4.2.3. Canadá²³

La Oficina del Auditor General de Canadá (OAGC) es la institución encargada de supervisar y evaluar el gasto público. Se trata de una entidad vinculada al Parlamento cuyo máximo responsable, el Auditor General de Canadá, es nombrado por un periodo no renovable de 10 años mediante una resolución conjunta de la Cámara de los Comunes y el Senado. Es una entidad plenamente independiente del Gobierno, que rinde cuentas exclusivamente al Parlamento. Su actividad se centra en el control de las actividades presupuestarias de los departamentos federales y de las agencias gubernamentales, la mayoría de las empresas públicas, así como otras organizaciones dependientes directa o indirectamente del Gobierno. De acuerdo con la estructura política federal de Canadá, la OAGC extiende su actividad a algunos gobiernos regionales, dependiendo en estos casos de sus respectivas asambleas legislativas.

De acuerdo con sus estatutos, la OAGC tiene la misión de proveer al Parlamento de información objetiva sobre los programas y actividades del Gobierno, basada en evidencias empíricas y obtenida a través de la realización de auditorías. Además, la OAGC ofrece asesoramiento profesional sobre estas cuestiones. Sus informes sirven para que los parlamentarios supervisen la gestión de los fondos públicos realizada

²³ https://www.oag-bvq.gc.ca/internet/english/admin_e_41.html

por el Gobierno en el desarrollo de las actividades públicas. Desde los años noventa del siglo XX, la OAGC cuenta con un Comisionado de Medio Ambiente y Desarrollo Sostenible, con responsabilidades de Auditor General, encargado de la supervisión de estas materias en los departamentos federales.

Por lo que respecta a los recursos humanos, para la incorporación de personal a la OAGC se sigue un proceso abierto de selección orientado a contratar a expertos de elevada cualificación, ofreciendo atractivas carreras profesionales. La OAGC también realiza, de forma regular, procesos formativos de especialización y actualización técnica, muy valorados entre los profesionales.

Por lo que respecta a su actividad, la OAG desarrolla tres tipos de auditorías:

- Auditorías financieras, dirigidas a supervisar los estados financieros y presupuestarios del gobierno central y de los gobiernos territoriales, así como de las empresas públicas y del resto de organizaciones dependientes del poder ejecutivo. Anualmente, la OAGC audita las cuentas públicas del Gobierno central y de los gobiernos regionales bajo su mandato. Estas auditorías son similares a las realizadas en el sector privado, pero también incluyen una supervisión del cumplimiento normativo que afecta a las actividades realizadas, así como de otros extremos que se consideren relevantes para los intereses del Parlamento. Estas auditorías suelen incluir el análisis del cumplimiento de objetivos y la evaluación de la fiabilidad de los sistemas de control financiero de los departamentos y agencias. La OAGC realiza anualmente alrededor de un centenar de auditorías financieras, destinando a las mismas alrededor de la mitad de sus recursos.
- Auditorías de gestión, cuyo contenido principal está orientado a evaluar si los programas de gasto público son ejecutados de acuerdo con los principios de eficiencia y eficacia, además de su cumplimiento en relación con el medioambiente. Estas auditorías también incluyen una valoración de los sistemas de seguimiento de objetivos implantados en los centros gestores para poder medir el impacto y la efectividad de los programas gestionados. Su realización se lleva a cabo con plena independencia del poder político, con el objetivo de ofrecer una evaluación objetiva y sistemática de la gestión de las actividades públicas, bajo la perspectiva del principio de *value for money*.²⁴

²⁴ Este principio trata de reflejar el compromiso, en términos de buen uso, asumido con los ciudadanos en cuanto a la provisión de bienes y servicios, como contraprestación de los recursos fiscales puestos a su disposición para tal fin.

Para desarrollar estas auditorías se realizan previamente exámenes profundos de los procedimientos de gestión, de la aplicación de mecanismos de control interno y del funcionamiento de los sistemas de información, de acuerdo con los criterios establecidos por la propia administración y de las mejores prácticas.

Estas auditorías, planificadas de acuerdo con las normas y protocolos establecidos, se llevan a cabo por profesionales de la evaluación especializados en áreas concretas de actividad. El Auditor General de Canadá dispone, por mandato legal, de una gran discrecionalidad para seleccionar las áreas de gobierno evaluadas. Esta capacidad de selección resulta muy importante, eligiéndose aquellas actividades que se consideran de interés prioritario para el Parlamento, tras un proceso previo de planificación de auditorías que incluye el análisis de riesgos (p.e. programas de salud pública, seguridad o de transporte), así como los recursos humanos y financieros disponibles, teniendo en cuenta que el plan puede extenderse hasta año y medio (anualmente se realizan entre 25 y 30 auditorías de gestión). La OAG suele recibir cada año un elevado número de propuestas de actividades específicas a auditar por parte de parlamentarios, miembros del Gobierno, incluso por ciudadanos particulares, siendo consideradas preferentemente las procedentes de las comisiones parlamentarias.

Por lo que respecta a la delimitación de las actividades a auditar, esta es muy variable y va desde programas muy concretos hasta áreas de actividad amplias, con participación de varios departamentos o agencias, incluyendo actividades que afectan transversalmente a la gestión pública (p.e. la seguridad en el uso de las tecnologías de la información, auditorías que afecten a resultados globales de la actuación gubernamental o del funcionamiento de la gestión financiera federal).

Los resultados de estas auditorías de gestión son incorporados a un documento en el que se contienen, de forma objetiva, las valoraciones positivas y negativas derivadas de los análisis realizados, así como las recomendaciones oportunas para mejorar el rendimiento de los centros y programas auditados. Su difusión se lleva a cabo mediante dos informes anuales –uno en primavera y otro en otoño– que el Auditor General remite al Parlamento para su conocimiento y discusión en las comisiones parlamentarias de Cuentas Públicas y, en su caso, de Medio Ambiente y Desarrollo Sostenible. En estas discusiones participan tanto gestores de las unidades evaluadas como los auditores de la OAGC. Tras

las mismas, se presenta un informe final a la Cámara de los Comunes, incorporando las recomendaciones al Gobierno. Con este procedimiento se materializa la rendición de cuentas del Gobierno ante el Parlamento, además de servir como impulso de mejoras de la gestión pública. De forma previa a su presentación en la Cámara de los Comunes, el Auditor General (o en su caso el Comisionado de Medio Ambiente y Desarrollo Sostenible) ofrece una vista previa confidencial de los informes, con una sesión de debate, y posteriormente, de acuerdo con el compromiso de máxima transparencia, convoca una rueda de prensa para exponer los resultados de estos informes, a la que asisten los responsables de los equipos que han realizado las auditorías.

- Exámenes especiales (*special examinations*), consistentes en una variante de las auditorías de gestión realizadas en las empresas públicas estatales sometidas a la jurisdicción de la OAGC, de acuerdo con el mandato de la Financial Administration Act, que obliga a que al menos una vez cada diez años estas entidades se sometan a esta supervisión. El cometido fundamental de estos exámenes es comprobar en qué medida los sistemas de control interno de las entidades aseguran la preservación de sus activos y que sus recursos son gestionados bajo criterios de eficiencia y eficacia. En promedio, la OAGC realiza anualmente tres auditorías de este tipo. Los resultados obtenidos en cada una de ellas se presentan en un informe que es remitido directamente al consejo de administración de la empresa pública, si bien en casos excepcionales se remite también al ministro sectorial y al Parlamento. Posteriormente, el consejo de administración envía un informe con las medidas adoptadas al ministro sectorial y al presidente de la Secretaría del Treasury Board, órgano del Gobierno encargado del control de las finanzas públicas. Estos informes se hacen públicos a través de la página web de la OAGC y de las empresas auditadas.

4.2.4. Estados Unidos²⁵

La General Accounting Office (GAO) es la entidad pública encargada de evaluar el gasto público del Gobierno estadounidense. Se trata de una institución independiente del gobierno que rinde cuentas al Congreso de los Estados Unidos. Su máximo responsable, el Comptroller General of the United States, es elegido por el Presidente de la Nación entre una lista de candidatos propuestos por una comisión bicameral del

²⁵ Sitios web: <https://www.gao.gov>; <https://www.cbo.gov>; <https://www.whitehouse.gov/omb/>.

Congreso integrada por representantes de los partidos Demócrata y Republicano, y su mandato se extiende por un plazo de 15 años.

Según su carta fundacional, la función principal de la GAO es dar soporte profesional y apoyar al Congreso en el cumplimiento de sus responsabilidades constitucionales y contribuir a mejorar el rendimiento del Gobierno, así como asegurar la rendición de cuentas y la exigencia de responsabilidades. Para ello se señala que su actividad debe proporcionar información oportuna, objetiva, justa y equilibrada, basada en hechos, independiente de los partidos políticos y no ideológica.

Las actuaciones técnicas de la GAO están fundamentadas en valores como la responsabilidad (*accountability*), la integridad de los análisis y el rigor profesional. Estos principios se aseguran bajo el estricto cumplimiento de los estándares profesionales de revisión y una evaluación exhaustiva y rigurosa. En cuanto al desarrollo de auditorías, estas se ajustan a los Principios Fundamentales de Auditoría (nivel 3) de los International Standards of Supreme Audit Institutions. Las actividades de la GAO se realizan a solicitud de los Comités y Subcomités del Congreso, o como consecuencia del mandato de leyes que así lo demanden, o de informe de las comisiones del Congreso.

En concreto, el apoyo a la actividad supervisora del Congreso se realiza mediante:

- a) la realización de auditorías en las agencias federales para determinar si los fondos públicos se emplean de forma eficiente y eficaz;
- b) evaluar en qué medida los programas y políticas gestionadas por el Gobierno cumplen los objetivos fijados;
- c) realizar análisis de políticas públicas y proponer alternativas para su posible consideración por el Congreso;
- d) investigar las denuncias de actividades ilegales o indebidas;
- e) emitir resoluciones y dictámenes jurídicos sobre reglas aprobadas por las agencias e impugnaciones planteadas.

Desde una perspectiva de asesoría, la GAO asume la misión de aconsejar al Congreso y a los responsables de las agencias ejecutivas federales acerca de los métodos y prácticas conducentes a conseguir un gobierno más eficiente, efectivo, ético, justo y receptivo, con un importante impacto en términos de ahorro presupuestario a través de la mejora de las leyes y decisiones. Desde el punto de vista de la transparencia, la GAO permite acceder a través de su página web a todos los informes de evaluación y supervisión realizados a las agencias ejecutivas y oficinas federales, así como a los estudios y recomendaciones de reforma propuestos²⁶.

²⁶ Estos informes se encuentran disponibles en http://www.gao.gov/key_issues/overview#t=2, además de un completo portal para investigadores <http://www.gao.gov/resources/researchers/overview#t=1>.

La Congressional Budget Office (CBO) es otra institución muy relevante en Estados Unidos en materia de evaluación económica de las actividades públicas. La CBO es una agencia federal creada en 1975 y ubicada dentro del poder legislativo norteamericano cuya función es proveer de información presupuestaria y económica al Congreso. La CBO se estructura en ocho divisiones: análisis presupuestario, análisis financiero, análisis de sanidad, pensiones y de largo plazo, análisis macroeconómico, estudios microeconómicos, seguridad nacional, análisis tributario y un servicio de gestión e información. Entre sus actuaciones más relevantes en materia de evaluación destaca el análisis económico de todos los proyectos legislativos, de forma bastante anticipada al inicio de su discusión parlamentaria. En determinados casos, los informes elaborados para los nuevos proyectos de gasto impiden su aprobación por incumplimiento de los objetivos de déficit público. La CBO también se encarga de analizar económicamente el proyecto anual del presupuesto federal. Sus estudios microeconómicos de eficiencia productiva por sectores de actividad y la realización de evaluaciones de impacto sobre programas aprobados son rigurosos y muy valorados.

En relación con la aplicación del ACB, su consideración como requisito para aprobar determinados programas de gasto público, generalmente de importe presupuestario elevado, es cada vez más habitual en algunos estados del país. En un reciente informe, las fundaciones The Pew Charitable Trusts y The John D. and Catherine T. MacArthur Foundation (PEW-MacArthur, 2013) concluyen que 50 estados y el Distrito de Columbia desarrollaron, entre 2008 y 2011, 348 estudios de ACB, si bien su contenido mostraba una amplia variedad. Una interesante conclusión de este estudio es que los estados donde el uso del ACB cuenta con un mayor arraigo institucional son también aquellos en los que se han adoptado reformas presupuestarias con mayor éxito en cuanto al uso más eficiente de los recursos públicos. No obstante, se advierte que el compromiso institucional con este tipo de análisis solo puede considerarse elevado en 11 de los 50 estados y en el Distrito de Columbia. El resto de estados no muestran un compromiso pleno con la adopción sistemática de estos métodos de evaluación rigurosa de programas y proyectos públicos, si bien el estudio revela una evolución creciente en su uso.

En 2019, la herramienta de análisis First Cost-Benefit Model de la PEW-MacArthur (2019) estimó para diversos gobiernos estatales y locales el retorno esperado de la inversión pública (ROI) para una serie de programas en funcionamiento y otros que estaban previstos poner en marcha. Los programas analizados correspondían, fundamentalmente, a las áreas de salud mental para adultos, protección infantil y

justicia juvenil. Sus resultados son considerados como una herramienta de gran utilidad, que permite a los responsables de la gestión de esos programas valorar sus beneficios y costes, con el fin de modificar sus planes de financiación, mejorando así la rentabilidad social de los recursos públicos.

Finalmente debemos destacar que, desde 2018, la Office of Management and Budget, de la Presidencia de los Estados Unidos, incorporó una oficina denominada Evidence and Evaluation Team, dependiente de la Office of Management and Budget (2024), cuya misión es impulsar la evaluación de políticas públicas y coordinar su desarrollo dentro de la administración federal estadounidense. No obstante, los departamentos ministeriales federales tienen autonomía para aprobar sus propios planes de evaluación.

4.2.5. Francia²⁷

En la última década Francia ha acometido importantes cambios en materia de evaluación de políticas públicas dentro de su iniciativa de Modernización de la Acción Pública. La Entidad Fiscalizadora Superior de la actividad del sector público francés es el Tribunal de Cuentas (Cour de Comptes), que cuenta con una antigua tradición en materia de fiscalización de la actividad pública. Además, existen 17 Tribunales de Cuentas de ámbito regional (departamental). En 2020, la nueva presidencia del Tribunal de Cuentas puso en marcha un plan estratégico denominado Jurisdicciones Financieras 2025 en el que se abordan tres grandes objetivos: a) servir mejor a los ciudadanos con informes más diversificados, oportunos y accesibles; b) reforzar y actualizar la actividad de evaluación de las políticas públicas económicas, sociales y medioambientales del país; c) promover una gobernanza más integrada y ágil a través del análisis de las evaluaciones de las políticas públicas y del asesoramiento a los centros responsables de la gestión pública.

Con anterioridad, en 2009 se había creado el Comité de Evaluación y Control para la coordinación de actividades de evaluación. La evaluación de políticas se concibe desde la coordinación con los organismos supervisados, en los que los departamentos ministeriales están obligados a establecer sus propios planes de trabajo. En el modelo institucional, desde 2012, dentro de la iniciativa de Modernización de la Acción Pública, el Gobierno propone un programa de evaluación. En 2018 se estableció un plan para evaluar el impacto de los proyectos legislativos en materias relacionadas con las políticas públicas sociales.

²⁷ <https://www.ccomptes.fr/fr>

En cuanto a recursos, el Tribunal de Cuentas francés cuenta con aproximadamente 1.500 empleados, de todos los niveles. En los últimos años se ha apostado por la incorporación de especialistas en evaluación. No obstante, más de la mitad de las evaluaciones son realizadas, mediante contratación externa, por consultoras privadas y por centros universitarios especializados. En 2022, dentro del Plan Estratégico 2025, se ha firmado un convenio de colaboración entre la Cour de Comptes y la Société Française d'Évaluation²⁸ con el fin de reforzar la formación especializada y el desarrollo de metodologías de evaluación basadas en estándares homologados a nivel internacional. El Centro de Acceso Seguro a los Datos (Centre d'Accès Sécurisé aux Données), creado en 2010 y en el que están involucrados varios ministerios y centros de investigación, se ocupa de la accesibilidad a la información precisa para los análisis de evaluación.²⁹

4.2.6. Países Bajos³⁰

La Ley de Cuentas del Estado en Países Bajos dispone que todos los proyectos de políticas públicas incluyan con claridad sus objetivos, así como la definición de los instrumentos que se aplicarán para su consecución. Además, en proyectos de infraestructuras, la evaluación ex ante es obligatoria. La evaluación de políticas públicas en Países Bajos es responsabilidad del Tribunal de Cuentas, institución que, además de encargarse de la supervisión financiera del sector público, analiza también la eficacia y la eficiencia de las políticas públicas. El Tribunal de Cuentas rinde sus informes al Parlamento.

Además, desde 2006 los ministros son responsables de realizar *policy reviews* que abarcan los programas públicos de períodos entre cuatro y siete años. Para ello, cada departamento ministerial posee una unidad de análisis del comportamiento presupuestario. En cambio, las evaluaciones son en su mayor parte desarrolladas, mediante contratación externa, por consultoras, centros universitarios especializados e institutos de investigación.

²⁸ <https://www.sfe-asso.fr>

²⁹ <https://www.casd.eu>

³⁰ <https://english.rekenkamer.nl>

4.2.7. Reino Unido³¹

En el Reino Unido la evaluación de los programas de gasto público la desarrolla fundamentalmente la National Accounting Office (NAO). La NAO evalúa el gasto público por mandato del Parlamento y es plenamente independiente del Gobierno. Su responsable, el Comptroller y Auditor General, es funcionario de la Cámara de los Comunes. En su estatuto, la actividad evaluadora de la NAO es considerada como esencial para hacer realidad la exigencia de responsabilidades (*accountability*) en la acción de los gobiernos en relación con la gestión del gasto público y la mejora de la calidad de los servicios públicos.

La NAO tiene una plantilla aproximada de 860 empleados. Entre sus mandatos está certificar las cuentas de los departamentos de la Administración Central, así como del resto de agencias y organismos gubernamentales. Como consta en sus estatutos, su función primordial es asegurar que los fondos recibidos por las unidades que gestionan los programas públicos se han utilizado de manera eficiente y efectiva, debiendo hacerse constar el cumplimiento de estos criterios económicos en los informes enviados al Parlamento. Sus informes son el resultado de las denominadas auditorías *value for money* realizadas para los diversos programas de gasto público, tanto en el ámbito central como local. Además, la NAO realiza informes de "recomendación y buenas prácticas" en los que se analizan los problemas organizativos y de funcionamiento económico-financiero de las diversas unidades gestoras, con el fin de mejorar la eficiencia en su comportamiento.

En relación con la mejora de la gestión económica del gasto público, el Gobierno británico creó en 2010 el Efficiency and Reform Group (ERG) dentro del Cabinet Office (Oficina del Primer Ministro). La principal función que se le asignó ha sido ayudar a los departamentos de gasto a lograr ahorros presupuestarios a través de mejoras de su eficiencia en la provisión de bienes y servicios, fomentadas desde el reforzamiento de la supervisión centralizada. En su funcionamiento participa también el HM Treasury (Ministerio de Hacienda), aunque cada ministerio es responsable de planificar y ejecutar sus propias evaluaciones. La actividad desempeñada por el ERG ha sido considerada clave en el proceso de consolidación de las finanzas públicas británicas acometido en los últimos años. A partir de los análisis del gasto público realizados por el ERG, el Committee of Public Accounts estimó unos ahorros de miles de millones de euros anuales por la supresión de ineficiencias, la reforma de los

³¹ Sitios web: <https://www.nao.org.uk>; <https://www.gov.uk/government/organisations/civil-service-government-social-research-profession>

procesos de gasto y los recortes aplicados en políticas consideradas de baja prioridad, si bien estos últimos se han visto condicionados seriamente por el mal comportamiento de la economía británica como consecuencia de la crisis de la pandemia de la Covid-19 y del Brexit. Los informes realizados por el ERG se centraban fundamentalmente en el análisis de costes de los centros gestores de los programas de gasto. No obstante, una de las críticas más frecuentes realizadas en estos informes tiene que ver con la falta de identificación clara de los objetivos a alcanzar por parte de las unidades de gasto, en muchos casos por una información deficiente de sus servicios de control interno, lo que a su vez dificultaba la realización de análisis de eficiencia comparativa entre unidades. En 2024, la ERG ha sido disuelta en el plan de reorganización administrativa del Reino Unido sin que se haya establecido qué unidad del Cabinet Office deberá asumir sus funciones.

En relación con los recursos humanos dedicados a la evaluación, hay que destacar el papel del Government Social Research Service (GSRS), organismo que en los últimos años ha sido clave para la extensión de la evaluación basada en el análisis cuantitativo y en la evidencia empírica. El GSRS está integrado por miembros relevantes del Gobierno y una red pluridisciplinar de investigadores. El Magenta Book (HM Treasury, 2020), impulsado por el GSRS, es la guía básica para evaluadores en el Reino Unido. Entre las funciones del GSRS está realizar investigaciones y asesorar a los gestores públicos en el ámbito social y del comportamiento, adoptando un enfoque científico, con el fin de aportar experiencia en enfoques aplicados de diseño y análisis cuantitativos y cualitativos de políticas públicas, la práctica de la evaluación, así como los procesos de toma de decisión para resolver problemas complejos. En la actualidad, el GSRS cuenta con más de 2.400 profesionales acreditados en evaluación, repartidos en más de 50 departamentos y oficinas públicas. Además, el GSRS desarrolla una labor fundamental en el mantenimiento y fomento de las habilidades profesionales, conocimientos y experiencia en materia de investigación social.

4.2.8. Suecia³²

En los años sesenta del siglo XX, Suecia fue pionera en la realización, de forma regular, de auditorías de resultados en relación con el desempeño del sector público (*performance audits*) y con su rendimiento social (*value for money audits*). Además, contó con la primera experiencia en la institucionalización de la evaluación de programas de gasto público desde un enfoque económico, con el propósito de valorar

³² Sitios web: <https://www.riksrevisionen.se/en/about-the-swedish-nao.html>; <https://www.jo.se/en/>

los resultados de la actuación pública desde criterios de eficiencia y eficacia, yendo más allá de la concepción tradicional de la auditoría financiera.

La agencia encargada de supervisar las finanzas públicas estatales es la Oficina Nacional de Auditoría de Suecia (Riksrevisionen, SNAO), cuyos antecedentes se remontan dos siglos atrás. Fue refundada en 1983, siendo continuadora de la actividad de la Agencia Nacional de Auditoría (Riksrevisionsverket), que desde los inicios de los años setenta ya venía realizando auditorías de gestión *value for money*. Se trata de una agencia independiente del Gobierno que de acuerdo con un mandato constitucional rinde cuentas exclusivamente al Parlamento (Riksdag). Su ámbito de actuación incluye la totalidad de las agencias gubernamentales, las empresas públicas estatales y los organismos del Gobierno, además de auditar el Informe Presupuestario Anual del Estado. La misión encomendada a la SNAO consiste en evaluar las actividades del Gobierno con el fin fundamental de contribuir a conseguir una gestión más eficaz por parte de la Administración del Estado. La SNAO está dirigida por tres auditores generales (*riksrevisorer*) designados por el Parlamento entre profesionales de reconocido prestigio y con mandatos improrrogables de siete años. Estos auditores generales tienen plena autonomía para establecer los planes de revisión y su alcance, así como el procedimiento de realización y los contenidos y recomendaciones a incluir en sus informes.

La SNAO define estatutariamente sus campos de actuación mediante el desarrollo de las siguientes actividades:

- Realización de auditorías financieras en las que se evalúan los estados financieros de las entidades gubernamentales, determinando su credibilidad y veracidad, así como si las cuentas presentadas son correctas y han sido obtenidas bajo el cumplimiento de las normas vigentes. Se realizan anualmente unas 250 auditorías financieras.
- Realización de auditorías de gestión (*performance audit*) dirigidas a evaluar la eficiencia de la actuación del sector público estatal. Este tipo de auditorías incluyen el análisis del cumplimiento de los objetivos de los programas presupuestarios, así como la valoración económica de la adecuación de la organización, las operaciones y procedimientos seguidos en las unidades gestoras para alcanzar dichos objetivos. Estas auditorías se extienden a la evaluación de impacto de los proyectos con financiación presupuestaria en prácticamente todas las áreas de actuación pública.

- La SNAO participa en la cooperación internacional para la mejora del control y evaluación de la actividad pública, con una especial dedicación al desarrollo de instituciones de auditoría y control en países en desarrollo, por mandato del Parlamento sueco, fomentando valores como la independencia y la profesionalidad de estas tareas. Además, la SNAO tiene una gran tradición internacional en la elaboración de metodologías de auditoría y evaluación, participando y organizando foros internacionales.

Los informes de las auditorías financieras realizadas por la SNAO son remitidos tanto al Gobierno sueco como a las entidades auditadas y, en su caso, a los departamentos ministeriales a los que se encuentran adscritas (aunque el modelo organizativo-institucional concede una muy elevada autonomía de actuación a las agencias públicas). Por lo que respecta a las auditorías de gestión, cada año la SNAO publica alrededor de 30 informes que son remitidos al Parlamento, el cual posteriormente solicita del Gobierno la emisión, en un plazo de cuatro meses, de un informe de valoración de los mismos que incluya las medidas correctoras adoptadas y las que tiene previsto tomar. Una vez valorado, el informe es remitido al Parlamento para que adopte las decisiones que se estimen oportunas.

Cada año, la SNAO elabora un Informe Anual de los Auditores Generales con una síntesis de las actuaciones más relevantes llevadas a cabo tanto en materia de auditoría financiera como de auditoría de gestión, el cual es remitido al Parlamento para su discusión y aprobación. Además, anualmente se emite un Informe de Seguimiento en el que se valoran las medidas adoptadas por el Gobierno y las entidades auditadas en relación con las recomendaciones y sugerencias incluidas en los informes de auditoría. La SNAO realiza también un informe anual de autoevaluación, que remite al Parlamento, en el que incluye un análisis detallado del cumplimiento de los objetivos establecidos para las divisiones de la institución, así como un análisis de su gestión económico-financiera. La política de transparencia de la SNAO es considerada un paradigma. Como se recoge en sus estatutos, los informes y sus metodologías se encuentran a disposición de la sociedad con total accesibilidad. Los informes realizados entre el año 2003 y 2023 se recogen en su sitio web³³.

Además de las actuaciones de la SNAO, en Suecia existe una larga tradición en la realización de evaluaciones de funcionamiento de los servicios públicos por parte de la Oficina Parlamentaria del Defensor del Pueblo. Esta realiza informes ocasionales sobre entidades y servicios públicos concretos, además de análisis de la regulación

³³ <http://www.riksrevisionen.se/en/Start/Audit-reports/>.

en diferentes sectores económicos. La calidad técnica de sus análisis suele ser alta y, en muchos casos, son un referente para la adopción de reformas de gestión de las políticas públicas, así como en las propuestas de reformas por parte del poder legislativo.

5. La institucionalización de la evaluación de políticas públicas en España: Valoración de la situación actual

La evaluación de las políticas públicas constituye un instrumento imprescindible para asegurar su eficacia y eficiencia, mandato recogido en el artículo 31 de la Constitución Española de 1978. Como se advierte en De la Fuente *et al.* (2021), su implantación, concebida como práctica sistemática y generalizada, lleva recorriendo un proceso largo y gradual en el que a fecha de hoy existe un amplio consenso en que aún se encuentra lejos de alcanzar unos niveles equiparables con los de países con mejor desempeño en este ámbito. Una de las razones ha sido, seguramente, la ausencia de una cultura asentada de la evaluación, tanto en las administraciones públicas como en la sociedad civil. No obstante, en los últimos años se ha abierto camino una mayor preocupación por el tema de la evaluación, en buena medida impulsada por las recomendaciones de la Comisión Europea, la OCDE y algunos otros organismos internacionales, casi siempre asociada con la necesidad de mejorar la asignación de recursos presupuestarios en un horizonte fiscal donde el cumplimiento de las reglas fiscales va a condicionar la evolución de nuestras finanzas públicas.

Más allá de las preocupaciones por la disciplina fiscal, posiblemente el principal activador para las reformas relacionadas con la institucionalización de la evaluación, la importancia de acertar en el diseño de las políticas públicas en un momento crucial, marcado por las consecuencias de la crisis originada por la pandemia de la Covid-19 y las tensiones inflacionistas vividas desde comienzos de 2022, aparece como otro elemento que puede ayudar a comprender esta preocupación.

Como hemos argumentado, hacer compatibles unas finanzas públicas de calidad, que den respuestas satisfactorias a los grandes retos a los que nos estamos enfrentando y que se extenderán durante bastantes años, como la transición ecológica, la transformación digital o el envejecimiento de la población, con un crecimiento económico robusto e inclusivo, exige aumentar significativamente la eficiencia del gasto público. Y eso solamente es posible de conseguir con su evaluación sistemática y rigurosa, en los términos que hemos venido exponiendo, y donde el adecuado funcionamiento del marco institucional en el que esta debe desarrollarse constituye un factor crucial.

El reconocimiento de esta realidad ha hecho que el Plan de Recuperación, Transformación y Resiliencia (Gobierno de España, 2021) incorpore algunas

propuestas para generalizar y reforzar los procesos de evaluación de las políticas públicas en España. Sin embargo, como veremos más adelante, su materialización, fundamentalmente a través de la Ley 27/2022, de 20 de diciembre, de institucionalización de la evaluación de políticas públicas en la Administración General del Estado, presenta importantes carencias, si se compara con las mejores prácticas internacionales, lo que hace albergar serias dudas sobre el alcance efectivo del marco aprobado. En esta sección concretamos estas valoraciones y planteamos las líneas de actuación que, en nuestra opinión, deberían conformar una reforma que sea efectiva, superando las limitaciones apuntadas.

5.1. Los antecedentes de la evaluación de políticas públicas en España

Con la aprobación de la Constitución Española, su artículo 31.2 introdujo en nuestro ordenamiento financiero y presupuestario la exigencia de que “la programación y ejecución del gasto público respondan a criterios de eficiencia y economía”. Desde entonces, con más decisión formal que efectiva, el legislador ha ido incorporando medidas normativas que afectan a la gestión presupuestaria de las políticas de gasto, así como la creación de algunos órganos administrativos con competencias en materia de control y evaluación, aunque, como veremos, en la mayor parte de los casos, con escasa efectividad en sus actuaciones.

La Ley General Presupuestaria (Ley 47, 2003, de 26 de noviembre) en su artículo 72 dispone que “El Ministerio de Hacienda, en colaboración con los distintos centros gestores de gastos, impulsará y coordinará la evaluación continuada de las políticas de gasto con la finalidad de asegurar que las mismas alcancen sus objetivos estratégicos y el impacto socioeconómico que pretenden”. Asimismo, la Ley 40/2015, de 1 de octubre, de Régimen Jurídico del Sector Público, en su artículo 3 incluye los principios generales que han de regir las actuaciones de las administraciones, entre los que, relacionados con el tema que nos ocupa, destacan: g) Planificación y dirección por objetivos y control de la gestión y evaluación de los resultados de las políticas públicas; h) Eficacia en el cumplimiento de los objetivos fijados; i) Economía, suficiencia y adecuación estricta de los medios a los fines institucionales; j) Eficiencia en la asignación y utilización de los recursos públicos. Para ello, ambas leyes fijan, en el ámbito administrativo, las competencias respecto del cumplimiento de estos principios en la Intervención General de la Administración del Estado (IGAE), centro directivo del Ministerio de Hacienda, así como, en las inspecciones de servicios de los departamentos correspondientes.

5.1.1. La evaluación de políticas públicas en la Administración General del Estado

Como se señala en De la Fuente *et al.* (2021), en el ámbito de la Administración General del Estado se llegan a contar más de un centenar de unidades con funciones o competencias atribuidas sobre la evaluación, en su concepción más amplia, si bien son apenas una cuarta parte las que desempeñan, aunque de forma parcial, tareas que pudieran ser consideradas como parte de la actividad evaluadora, aunque mayoritariamente centradas en actividades de seguimiento, auditoría, supervisión y análisis financiero; otro tercio estaría dedicado a actividades relacionadas con la inspección y calidad de servicios y el funcionamiento de las organizaciones encargadas de su provisión; y el resto tendría encomendadas tareas de supervisión, esencialmente administrativa, de la aplicación de la normativa legal. En general, este panorama pone de manifiesto una gran dispersión de objetivos en esas tareas, falta de uniformidad y sistematización de la práctica de la evaluación, así como una disparidad de criterios que impiden delimitar con claridad lo que se pretende con la evaluación de políticas públicas. En la práctica, la actuación de la IGAE se ha venido centrando en el control contable, financiero y de legalidad, con unos estándares y procedimientos que no corresponden a lo que, en un sentido estricto, pueden identificarse con la evaluación de las políticas públicas.

De hecho, este diagnóstico aparece reconocido, de forma soslayada, en la exposición de motivos de la Ley de institucionalización de la evaluación de políticas públicas en la Administración General del Estado, cuando se reconoce que “en nuestro país, en los últimos años se han producido avances importantes en materia de evaluación, aunque aún no se dispone de un marco integrador común que proporcione las herramientas y elementos básicos para impulsar una política de evaluación de las políticas públicas”, para a continuación señalar que “las diferentes administraciones públicas vienen desarrollando experiencias de evaluación, si bien la mayoría tiene carácter sectorial” y concluir que el propósito de esa nueva ley -reconocida para la modernización de la Administración Pública como hito en el PRTR- es “implantar en la Administración General del Estado la evaluación sistemática de políticas públicas y la creación de un organismo con alto grado de autonomía, a partir del actual Instituto para la Evaluación de Políticas Públicas”.

En 2007 se produjo el que podemos considerar el primer intento ambicioso de sistematización, la creación de la Agencia Estatal de Evaluación de la Calidad de los Servicios y de las Políticas Públicas (AEVAL). Adscrita al Ministerio de Administraciones Públicas, la AEVAL comenzó a funcionar con la misión de promover

y realizar evaluaciones de las políticas y programas de la Administración Central con el fin de favorecer el uso racional de sus recursos y la mejora de la calidad de sus servicios. De acuerdo con esa dependencia orgánica, se establecía que su plan de trabajo anual, incluyendo los programas y políticas estatales, se aprobaría por el Consejo de Ministros. Es relevante señalar que, en materia de recursos humanos, AEVAL contaba inicialmente con los empleados públicos de la antigua Dirección General de Inspección, Evaluación y Calidad de los Servicios del Ministerio de Administraciones Públicas, lo que limitó desde el principio la posibilidad de contratación de personal especializado en técnicas avanzadas de evaluación.

Como señalan Albi y Onrubia (2018), el gran problema de diseño de la AEVAL fue, además del apuntado de la escasez de personal cualificado, la total falta de independencia del poder ejecutivo, con una dependencia orgánica directa de un departamento ministerial. Además, aunque su estatuto señala la "independencia de criterio, dictamen y juicio en la realización de sus trabajos" (artículo 7) como el primero de sus principios básicos de actuación, la AEVAL era dirigida por un presidente y unos consejeros nombrados por los Ministerios, en su mayoría altos cargos de estos, y como hemos dicho, su plan de trabajo debía ser aprobado por el Consejo de Ministros. Así, su actividad terminó concentrándose en la realización de estudios que revisaban la calidad de los servicios públicos, o en análisis de orientación económica generalista que no encajaban, de acuerdo con los estándares profesionales en este campo, con lo que puede considerarse evaluación de impacto de programas o proyectos públicos, o en análisis ex ante para el diseño de políticas.³⁴

Tras un tiempo sin apenas actividad, la AEVAL fue disuelta en 2017 por el Real Decreto 769/2017 que aprobaba una reorganización del Ministerio de Hacienda y Función Pública. Esta norma establecía que sus funciones serían asumidas por la Secretaría de Estado de Función Pública, a la vez que creaba un nuevo Instituto para la Evaluación de Políticas Públicas, con rango de subdirección general dentro de esta Secretaría de Estado, cuyas funciones serían "la evaluación de las políticas públicas, especialmente en los ámbitos educativo, sanitario y social a nivel nacional, así como ofrecer el apoyo instrumental necesario para realizar los análisis que requieran los procesos de modernización o planificación que se impulsen desde la Secretaría de Estado de Función Pública" (art 13.7). Desde aquel momento hasta la reciente aprobación de la Ley 27/2022, ese Instituto no realizó tampoco actividades de evaluación de las que se tuviera noticia.

³⁴ Recientemente, en la web de la Secretaría de Estado de Función Pública se ha dado acceso abierto a los informes de evaluación de políticas públicas realizados por la AEVAL hasta 2016.

La Comisión para la Reforma de las Administraciones públicas (CORA), creada por el Gobierno en octubre de 2012 para analizar la situación del sector público español y hacer propuestas para mejorar su funcionamiento y reducir sus costes, en su informe de la Comisión (CORA, 2013) impulsó la normativa relacionada con la evaluación ex ante y ex post en la Administración Pública del Estado, destacando la incorporación en la Ley 40/2015, de régimen jurídico del sector público, la obligación de elaborar, para cada norma con rango de ley o reglamento, una memoria de análisis del impacto normativo que cabría considerar, al menos potencialmente, como un informe de evaluación ex ante. A pesar de la regulación en el Real Decreto 931/2017 de un reglamento de procedimiento para desarrollar este tipo de análisis, la realidad desde entonces ha sido que dicha exigencia se ha convertido en un mero trámite documental cuyo contenido informativo suele ser de escaso interés para los fines inicialmente perseguidos.³⁵

Posiblemente, antes de la aprobación de la Ley 27/2022 de institucionalización de la evaluación de políticas públicas en la Administración General del Estado, impulsada como hito en el Plan de Recuperación, Transformación y Resiliencia, la iniciativa más destacada en el ámbito de la evaluación ha sido la puesta en marcha de las Spending Review realizadas por la AIREF y que ya fueron reseñadas en la segunda sección de este documento. Las Spending Review tienen por objeto mejorar la eficiencia y efectividad del gasto público a través de un escrutinio económico de programas concretos de gasto, si bien su alcance puede ser dispar como señalan Gambau y López Herrera (2023).

Al residenciar su realización en la AIREF, se generaba un impulso que por primera vez sacaba las competencias evaluadoras del ámbito de dependencia gubernamental, si bien es cierto que circunscritas al alcance de ese tipo concreto de revisión de programas de gasto. Aunque, sin duda con limitaciones funcionales, hay que decir que las Spending Review han contribuido positivamente a potenciar la cultura de la evaluación de políticas públicas en España, especialmente en lo referente a la importancia de la independencia de los órganos que deben llevarlas a cabo. También merece una valoración positiva el hecho de que en su realización se haya tenido en cuenta la necesidad de contar con profesionales especializados, si bien la escasez de efectivos de la AIREF, junto con las limitaciones para la incorporación de profesionales a su plantilla, han dado lugar a que una parte bastante importante de la actividad

³⁵ En la web de Transparencia se pueden consultar, no sin una complejidad no menor, los sucesivos planes anuales normativos y los correspondientes informes de evaluación, además de las memorias de leyes específicas. En ellos puede contrastarse la habitual falta de detalle y la carencia de contenido de interés.

evaluadora de las Spending Review haya sido externalizada, fundamentalmente con firmas de consultoría.

5.1.2. La evaluación de políticas públicas de las comunidades autónomas

Dado el gran nivel de descentralización de las políticas públicas en España, entre otras, las correspondientes a tres grandes pilares del Estado del bienestar, como son la sanidad, la educación y los servicios sociales, no se puede completar un panorama de la evaluación de políticas públicas sin considerar las experiencias de las comunidades autónomas en este terreno. El artículo de Aurelia Valiño Castro, "La evaluación económica en el análisis de las políticas públicas: el estado del arte en España", publicado en el volumen 103 de *Ekonomiaz*, ofrece una revisión de las iniciativas de evaluación de políticas públicas en el ámbito autonómico, esbozando sus logros y sus debilidades.

Comenzando por estas, en el análisis de Valiño (2023) se destaca la falta de coordinación existente entre la Administración central y las comunidades en lo referente al diseño de las políticas y la generación de información necesaria para su evaluación. Se resaltan, en esta línea, los intentos de la suprimida AEVAL por crear, mediante convenios bilaterales, un grupo de trabajo operativo, al que se adhirieron algunas comunidades (Baleares, Cataluña, Castilla-León, Extremadura, Navarra, País Vasco y Comunidad Valenciana), y al que se unió la Federación Española de Municipios y Provincias. Posteriormente, tras la supresión de la AEVAL, la AIReF sería la encargada de jugar un papel en la coordinación en materia de evaluación de políticas en el ámbito autonómico, si bien en este caso la fórmula elegida era la contratación por parte de los gobiernos autonómicos de proyectos de evaluación de políticas públicas sectoriales y programas concretos. En lo referente a los aspectos de institucionalización de la evaluación pública en las comunidades autónomas, solamente la comunidad de Castilla y León suscribió un encargo a este organismo.

Un rasgo común destacado en muchas comunidades es que la exigencia de información por parte de la Unión Europea sobre el seguimiento y evaluación de la gestión de fondos comunitarios para financiar políticas concretas, como sucede con las políticas de empleo, ha llevado a los gobiernos autonómicos a potenciar la realización de evaluaciones de seguimiento e impacto, así como de las evaluaciones de diseño de programas, un sucedáneo de evaluación ex ante. En este sentido, la gestión autonómica de los nuevos fondos de la Unión Europea del programa Next Generation EU y sus fondos vinculados, podría impulsar la práctica evaluadora y, por ende, extender la cultura de evaluación.

Desde el punto de vista más institucional, encontramos comunidades autónomas que han legislado sobre evaluación, creando en algunos casos instituciones de evaluación autonómicas donde varía el grado de independencia. Por regla general, se observa que las comunidades que han incorporado este tipo de instituciones habrían mostrado una mayor actividad evaluadora, tanto en número de evaluaciones realizadas como en su amplitud temática. Asimismo, estas instituciones han promovido la aprobación y difusión de guías de evaluación y ofertas formativas para el personal evaluador.

En lo referente a las experiencias institucionales autonómicas para el desarrollo de la evaluación pública, Valiño (2023) recoge un inventario bastante completo en la tabla nº 1 de su artículo³⁶. En él se presenta información detallada de las instituciones de evaluación de cada comunidad, allí donde existen, incorporando enlaces a sus respectivas páginas web, así como a las guías de evaluación disponibles. Finalmente, el inventario muestra las principales áreas sectoriales de políticas públicas en las que las distintas comunidades han desarrollado evaluaciones.

5.2. La evaluación de políticas en el Plan de Recuperación, Transformación y Resiliencia, y la Ley 27/2022, de institucionalización de la evaluación de políticas públicas en la Administración General del Estado

El Plan de Recuperación, Transformación y Resiliencia aprobado por el Gobierno de España (2021) y validado por la Unión Europea incluye varias propuestas para reforzar la evaluación de las políticas públicas, con tres líneas de actuación: aprobación de cambios normativos, creación de un nuevo organismo de evaluación, y la institucionalización de las Spending Review. Estas reformas se contemplan en el Componente 11 del Plan dedicado a la Modernización de las Administraciones Públicas.

Estas propuestas se han visto materializadas, como ya se adelantó, en la aprobación de la Ley 27/2022, de 20 de diciembre, de institucionalización de la evaluación de políticas públicas en la Administración General del Estado. Como se recoge al comienzo de su exposición de motivos, esta ley estaría atendiendo a la insistencia de la Unión Europea, en los últimos años, sobre la conveniencia de crear un marco regulador del proceso de evaluación de las políticas públicas, situándolo en el nivel central de Gobierno. Los principales contenidos de la ley se recogen en el siguiente cuadro, elaborado a partir de De la Fuente (2022).

³⁶ <https://www.euskadi.eus/web01-a2reveko/es/k86aEkonomiazWar/ekonomiaz/abrirArticulo?idpubl=100®istro=10>

Cuadro 1. Principales medidas incorporadas por la Ley 27/2022, de institucionalización de la evaluación de políticas públicas en la AGE

<p>• Organización institucional</p> <ul style="list-style-type: none"> – En cada ministerio se designará una unidad encargada de la coordinación y seguimiento de la actividad de evaluación de políticas (art. 27). – Se prevé la creación mediante decreto de la Comisión Superior de Evaluación como órgano de coordinación, cooperación y participación de los distintos ministerios en materia de evaluación de políticas en la AGE (art. 28). – Se prevé la creación del Consejo General de Evaluación como órgano consultivo, asesor y de participación de la sociedad civil en el ámbito de la evaluación de políticas públicas (art. 31). – Se autoriza la creación de la Agencia Estatal de Evaluación de Políticas Públicas, considerada como organismo público con personalidad jurídica propia adscrito a la Secretaría de Estado de Función Pública (art. 29). Le corresponderá la supervisión y coordinación del sistema estatal de evaluación de políticas públicas, el apoyo a los departamentos ministeriales y la realización de las evaluaciones que se determine. También será responsable, entre otras materias, de la formulación de metodologías de evaluación y la selección de los correspondientes indicadores de seguimiento y evaluación, la elaboración de la propuesta del Plan de Evaluaciones Estratégicas del Gobierno y la aprobación de los Planes de Evaluación Departamentales, del desarrollo de un mecanismo de seguimiento de los resultados de las evaluaciones. – Mientras se pone en marcha la Agencia Estatal de Evaluación de Políticas Públicas, el Instituto para la Evaluación de las Políticas Públicas asumirá sus funciones esenciales (DT única).
<p>• Políticas sujetas a evaluación</p> <ul style="list-style-type: none"> – Se detallarán en el Plan de Evaluaciones Estratégicas del Gobierno (art. 24), elaborado por la Agencia de Evaluación, que será aprobado cuatrienalmente por el Consejo de Ministros y en el Plan de Evaluación Departamental (art. 25), que será elaborado cada dos años por los distintos Ministerios y aprobado por la Agencia de Evaluación, previa consulta de la Comisión Superior de Evaluación. Cada ministerio será responsable de que se realice una evaluación ex ante y otra ex post de aquellas políticas de su competencia que tengan especial relevancia (art. 26).
<p>• ¿Quién es el responsable de hacer las evaluaciones?</p> <ul style="list-style-type: none"> – Las evaluaciones ex post de revisión del gasto público (Spending Review) serán responsabilidad de la Autoridad Independiente de Responsabilidad Fiscal, AIReF (art. 3.2). – La evaluación se realizará por un equipo externo al órgano responsable de la política a evaluar, aunque podrán constituirse equipos mixtos en los que participen responsables de la planificación (art.15.2).
<p>• ¿Qué se hace con las evaluaciones?</p> <ul style="list-style-type: none"> – Los informes de evaluación se publicarán, en el plazo máximo de un mes, en el portal de transparencia de la Administración General del Estado (Art. 22.4). – Tras la publicación del informe final de evaluación, en un plazo máximo de dos meses, el órgano responsable de la política evaluada deberá elaborar un “informe de resultados” indicando cómo se va a responder a las recomendaciones del informe de evaluación y explicando, en su caso, las razones por las que se ha optado por no aplicarlas (art. 23).
<p>• Análisis previo de evaluabilidad</p> <ul style="list-style-type: none"> – Se introduce la figura del “informe de evaluabilidad” referido a los documentos de planificación estratégica (art. 11).
<p>• Indicadores de seguimiento y evaluación de políticas públicas</p> <ul style="list-style-type: none"> – Se pretende que las políticas públicas cuenten desde su diseño con indicadores de seguimiento y evaluación ex ante y ex post (art. 12, art. 13). – Se pretende que estos indicadores se integren en un Sistema de Indicadores Comunes (art. 14).
<p>• Otras cuestiones</p> <ul style="list-style-type: none"> – La ley excluye explícitamente de su ámbito de aplicación, entre otras cosas, cualquier actividad de evaluación que sea objeto de regulación específica (art. 3.3).

Fuente: De la Fuente (2022).

En su artículo del número 103 de *Ekonomiaz*, De la Fuente y Novales (2023), así como Novales (2023), valoran de forma bastante negativa el contenido de la Ley 27/2022 por sus carencias, ambigüedades e imprecisiones. En concreto, se pueden identificar los siguientes problemas.

- En gran medida, la ley ha optado por recuperar la antigua figura AEVAL, con las limitaciones que ese modelo ya planteó, fundamentalmente las correspondientes a la dificultad de contar con personal especializado con los conocimientos técnicos necesarios, así como la falta de independencia de sus responsables y de su equipo directivo en relación con los organismos responsables de los programas a evaluar, lo que impide someter de forma efectiva las actuaciones políticas a un análisis riguroso.
- Se opta por ubicar la Agencia de Evaluación dentro de la Secretaría de Estado de Función Pública, si bien recientemente han surgido noticias sobre la posibilidad de su adscripción en el ámbito del Ministerio de Economía. A falta de conocer sus futuros estatutos, esto supone que la composición y designación de sus órganos de dirección queden fuertemente comprometidos en su independencia efectiva, tanto respecto del departamento ministerial de adscripción como del Gobierno en última instancia. Esto también se convertirá, previsiblemente, en un importante obstáculo para la incorporación de personal externo con la formación técnica requerida en materia de evaluación, que en España se concentra fundamentalmente en las universidades y centros de investigación.
- La norma legal no recoge ningún tipo de mejoras metodológicas ni de procedimiento técnico en relación con la práctica de la evaluación de las políticas públicas. En este sentido, la norma se limita a enunciar los elementos propios de la evaluación, pero de forma muy superficial.³⁷ Como señalan De la Fuente y Novales (2023), frente a la escasa concreción metodológica, en el texto aparecen referencias que resultan realmente extrañas respecto de los estándares académicos en el campo de la evaluación económica, que parecen confundir los objetivos de esta y los criterios de objetividad técnica bajo los que debería operar (p.e. las referencias a la evaluación como “mecanismo de lucha contra la inequidad” y el énfasis que se da a la “participación de los agentes implicados y de los sectores de la ciudadanía afectados”).

³⁷ Para una valoración más detallada, véase de la Fuente (2022).

- Es previsible que la ley genere una carga administrativa importante, sin que a priori parezca justificado. De hecho, como advierte Novales (2023:586), el afán extensivo de la norma –“realizar análisis de evaluabilidad de la totalidad de anteproyectos, programas, planes estratégicos y cualquier iniciativa que desarrolle políticas públicas, con el fin de verificar la calidad y precisión del diagnóstico realizado, la viabilidad, coherencia y consistencia del diseño elaborado, y la claridad, suficiencia y accesibilidad (art. 11)”-, se antoja una tarea ciertamente inasumible, tanto por el número de informes que deben realizarse como por su complejidad, tal y como se recoge detalladamente en la norma. Parecería más lógico y razonable que la norma hubiese establecido una sistemática temporal de realización de los planes de evaluación, según la función de cada tipo de evaluación y aplicable por áreas temáticas concretas, seleccionando los programas y proyectos de acuerdo con su relevancia social, así como por los volúmenes de recursos implicados. De esta manera, los planes de evaluación deberían establecer agendas de prioridades, algo fundamental para supervisar *a posteriori* su cumplimiento. Hay que decir que una dimensión de la actividad evaluadora como la que se contempla en la norma legal parece sencillamente desbordante y, hasta donde se conoce, no existe con tal amplitud en ningún país de referencia en esta materia. A nuestro juicio, esto será a medio plazo, previsiblemente, un factor determinante de la inacción, como ya ha sucedido en intentos anteriores de implantación de un marco integral de evaluación.
- En relación con los recursos humanos especializados, esenciales para el éxito de las evaluaciones, la ley simplemente se limita a señalar en su art. 15 que “Las evaluaciones las realizarán equipos externos a los órganos gestores de las políticas evaluadas”, sin aclarar, como señala Novales (2023), qué requisitos formativos deberán cumplir sus miembros, cómo serán los procesos de selección o qué órganos serán los responsables de su elección para garantizar la independencia requerida. Igualmente, tampoco existe mención a los procesos de actualización formativa en el campo de la evaluación, algo que la literatura especializada suele destacar como un factor clave en la especialización del personal dedicado a la evaluación.
- En materia de evaluación ex post, Novales (2023) valora la ley como escasamente ambiciosa. Destaca, a su vez, su poca claridad a la hora de diferenciar entre las alternativas disponibles para llevar a cabo las evaluaciones -de eficiencia, de efectividad, de impacto o las Spending Review realizadas por

la AIReF-, y la función que debe desempeñar cada una de ellas en el modelo institucional que pretende establecer la norma legal. Respecto al papel de cada tipo de evaluación, en ningún momento el texto legal plantea un marco de desarrollo de las actividades de evaluación, estructurado en línea con el contenido que hemos expuesto en las secciones iniciales de este documento.

- Otro aspecto criticable tiene que ver con la falta de previsión respecto de las necesidades estadísticas concretas que todo proceso de evaluación de políticas públicas aparea. Definir, en paralelo al diseño de la política, un procedimiento de recogida sistemática de la información proporcionada por su aplicación es un requisito esencial para desarrollar evaluaciones rigurosas. Se trata de un estándar aceptado en la literatura que requiere seleccionar anticipadamente los contenidos informativos que deben obtenerse por parte de las unidades encargadas de gestionar las políticas. A su vez, esto obliga a contar en los centros gestores con servicios de control interno operativos, con personal especializado, y que deben asumir la responsabilidad de generar esa información relevante para que puedan llevarse a cabo las evaluaciones.

En resumen, como destaca Novales (2023), el marco legal de institucionalización de la evaluación recientemente aprobado, a falta de la creación de la Agencia de Evaluación prevista, es una norma llena de ambigüedades. Su lectura no arroja luz sobre quién debe ejercer las responsabilidades técnicas ni institucionales de la evaluación, ni cuál es la estructura sistemática en la que se aplicarán los diferentes tipos de evaluación, de acuerdo con su funcionalidad específica, dentro del modelo integral con el que se quiere contar. En este sentido, la ley tampoco hace referencia concreta a las metodologías estándar de evaluación de políticas públicas a utilizar en cada caso, y sobre las que, como hemos dicho, existe una amplia literatura.

Asimismo, la norma legal no especifica de forma precisa quiénes deben ser los destinatarios de los informes de evaluación, ni las consecuencias, en términos de rendición de cuentas y de exigencia de responsabilidades, que de los mismos deberían derivarse; limitándose a indicar la conveniencia de incorporar a las políticas las mejoras de gestión recomendadas, pero sin establecer para ello un procedimiento institucional concreto.

6. Líneas para una reforma eficaz de la evaluación pública en España

Como se ha defendido a lo largo de este documento, la evaluación económica es inherente a la utilización adecuada de los recursos públicos, escasos y susceptibles de destinarse a asignaciones alternativas. En este sentido, los ahorros potenciales de recursos que puede proporcionar la evaluación a medio y largo plazo deberían ser razón suficiente para su práctica sistemática en el ejercicio de las actividades públicas. Pensamos que este argumento constituye, por sí mismo, un incentivo suficiente para una incorporación institucional efectiva al marco en el que se desarrolla la actividad financiera del sector público en relación con las políticas públicas.

Sin embargo, es frecuente que la evaluación se vea por parte de los políticos con responsabilidades de gobierno y de los gestores públicos más como una intromisión en su autonomía de decisión, que como un instrumento eficaz para generar información destinada a mejorar la calidad de las finanzas públicas y, consecuentemente, esencial para conseguir aumentos de bienestar social. En este sentido, la literatura sobre institucionalización de la evaluación de políticas públicas advierte de los comportamientos de cálculo estratégico que afectan tanto a la configuración de los programas y proyectos públicos, por un lado, como a los intereses individuales y corporativos, por otro, que suelen dificultar, cuando no bloquear, los avances reformadores conducentes a una implantación efectiva -no retórica- de la cultura de la evaluación. Como se apunta en Albi y Onrubia (2018), no debe olvidarse que llevar a cabo evaluaciones rigurosas de las actividades públicas no es en absoluto una tarea sencilla, ya que los requisitos a cumplir, como hemos reseñado en este documento, son muchos. Entre todos ellos, debemos destacar la necesidad de contar con especialistas en evaluación económica con formación adecuada en las técnicas a utilizar y que sean conocedores de los estándares presentes en las mejores prácticas en este campo a nivel comparado.

Si bien la formación y capacitación profesional de los especialistas que han de realizar las tareas de evaluación, junto con la disponibilidad de recursos económicos suficientes para llevarlas a cabo, son sin duda elementos indispensables para que la evaluación de las políticas y programas de gasto público pueda ser completada con las garantías de rigor necesarias, los requisitos de carácter institucional resultan igualmente imprescindibles. En este sentido, ocupa un papel preeminente la independencia efectiva de las instituciones encargadas de realizar las evaluaciones.

Es evidente que la dependencia orgánica, funcional o siquiera financiera entre el responsable de gestionar las políticas en cualquiera de sus niveles y la institución responsable de su evaluación constituye un escenario de conflicto de intereses no deseable, que limita seriamente el alcance de la práctica y, por tanto, la utilidad comentada.

Además, no podemos olvidar que para una práctica efectiva de la evaluación el modelo organizativo del sector público también aparece como un condicionante fundamental. Como se advierte en Onrubia (2005a, 2005b), sin sistemas de control interno capaces de generar información económica suficiente y de calidad sobre la actividad de los centros gestores y sin sistemas de presupuestación realmente vinculados a objetivos resulta muy complicado, por no decir imposible, poner en marcha procesos de evaluación rigurosos.

Hay que destacar también la importancia que tiene en la implantación de un modelo institucional de evaluación económica la voluntad política (*political will*, en su expresión en inglés). Como destacan Albi y Onrubia (2018), una gran parte de las decisiones relacionadas con el diseño del marco institucional de evaluación y su posterior funcionamiento se adoptan en la esfera política y necesitan el respaldo de la acción de gobierno. En este sentido, la implantación de una cultura de la evaluación económica y su valoración por parte de la sociedad condiciona a largo plazo esa voluntad política. La exigencia de responsabilidades hacia los responsables de la gestión pública, junto con la transparencia en la acción política vinculada a la misma, aparecen también como factores determinantes. De hecho, estos factores suelen estar muy arraigados en los países considerados con mejores prácticas en materia de evaluación pública.

Desde la perspectiva del control de la gestión pública, la institucionalización de la evaluación supone un cambio en la orientación de la supervisión del gasto público, desde la "lógica de los medios" hacia la "lógica de los resultados". Como se destaca en Onrubia (2005a), este cambio es algo esencial para conseguir el mejor aprovechamiento por parte de la sociedad de los recursos aportados por los ciudadanos a través del sistema fiscal. Se trata de la idea inspiradora del principio anglosajón de buena gestión conocido como *value for money*, inherente a los principios de rendición de cuentas y exigencia de responsabilidades (*accountability*).

Coincidimos con la valoración negativa que una amplia mayoría de analistas tienen sobre el estado de la evaluación económica de las actividades públicas en España.

Como señalan Albi y Onrubia (2018), tras más de cuatro décadas desde la aprobación de la Constitución de 1978, la evaluación puede considerarse en España como una asignatura pendiente. Después de algunos intentos voluntaristas como el comentado de la creación de la AEVAL, las limitaciones y deficiencias tanto técnicas como institucionales condicionan la realización de evaluaciones de acuerdo con los estándares y mejores prácticas recogidos por la literatura considerada como de referencia. Hay que decir que, posiblemente, la concepción extremadamente procedimentalista de la gestión pública española es un obstáculo que viene condicionando la implantación de una cultura de la evaluación, ya que la práctica de los controles de legalidad y regularidad contable acaban arrinconando a la supervisión de naturaleza económica.

De las experiencias de evaluación pública desarrolladas en el ámbito autonómico debemos destacar que, a diferencia de lo observado en el caso de la Administración General de Estado, un buen número de comunidades autónomas, aunque con limitaciones, han llevado a cabo esfuerzos importantes por desarrollar marcos institucionales propios en materia de evaluación. La creación de instituciones de evaluación, aunque con diverso grado de independencia, ha servido para potenciar una incipiente cultura de evaluación, con aprobaciones de guías metodológicas y programas formativos en la materia. Estas iniciativas se han visto potenciadas, a su vez, por la gestión autonómica de fondos de la Unión Europea, lo que ha impuesto la práctica de actividades de evaluación tanto del impacto de los programas financiados como de su diseño y seguimiento. No obstante, por regla general no se cuenta con planes de coordinación ni de compartición de información entre las diferentes comunidades ni de estas con la administración central, ni instituciones como la AIReF.

Esta orientación hace que el control interno encargado de proporcionar información de calidad para la evaluación –que abarca la determinación de objetivos e indicadores de actividad, la contabilidad de costes, los sistemas de seguimiento de objetivos, etc.– se encuentre anclado en principios de gestión propios de un sistema de administración presupuestaria de mediados del siglo XX. De hecho, estas carencias hacen que el control interno dentro de los centros gestores de los programas de gasto se dedique, en gran medida, a la preservación de los medios puestos a disposición de los gestores, olvidando casi por completo su papel como fuente generadora de información necesaria para la evaluación económica (Onrubia, 2005a). En relación con este papel, la existencia de un modelo de control interno exclusivamente de carácter horizontal, como el realizado por los órganos de Intervención, no fomenta la iniciativa ni la responsabilidad de los gestores en la generación de esa información.

De hecho, en los países considerados de referencia, las agencias y centros responsables de gasto suelen contar con fuertes servicios de control interno, bajo la responsabilidad de los gestores ejecutivos, si bien suelen estar bajo la supervisión técnica de órganos horizontales de control interno.

También debemos reconocer la falta de encaje de la evaluación en el marco institucional de la supervisión del gasto público establecido por la Constitución Española de 1978. La larga existencia del Tribunal de Cuentas y su reconocimiento constitucional como supremo órgano fiscalizador de la actividad económico-financiera del sector público llevó, en principio, a algunos especialistas a considerar a este como una institución superior de control independiente del poder ejecutivo, homologable a las existentes en los países con una mejor tradición en materia de evaluación. Sin embargo, como señalan Albi y Onrubia (2018), sus actividades de control, en relación con los principios de eficiencia, eficacia y economía, se encuentran prácticamente arrinconadas, desplazadas por la dedicación de sus recursos a la realización, prácticamente en exclusiva, de tareas de fiscalización centradas en los aspectos de legalidad y regularidad contable. Por supuesto, en esta orientación tiene mucho que ver la falta de adecuación de los nombramientos de sus máximos responsables, bastante alejada de los estándares formativos requeridos por la práctica profesional en el campo de la evaluación económica. La dispersión de objetivos, la insuficiencia crónica de recursos, junto con una estructura organizativa bastante anticuada, han trasladado a la sociedad una percepción de inoperancia en materia de evaluación económica. De hecho, resulta chocante que, a pesar del papel que le da la Constitución, la reciente ley de institucionalización de la evaluación de políticas públicas parece obviar su existencia, dejándola de facto fuera de ese marco institucional elegido.

A la vista de este diagnóstico poco alentador, nos parece inaplazable abordar un proceso de reflexión acerca del papel de la evaluación económica de las políticas públicas, que se extienda a la reforma de las instituciones que deben llevarla a cabo, tanto al ámbito del poder ejecutivo como del parlamentario, por supuesto, con la pertinente coordinación. Este proceso debe, a nuestro juicio, arrancar de un diagnóstico integral del funcionamiento de la gestión pública, incluida su conexión con los aspectos de la gestión y el control presupuestario. En este sentido, pensamos que estamos ante una ausencia de modelo de control y supervisión, en el que se establezca un reparto claro de las competencias en la evaluación, atendiendo a las diversas facetas, tanto temporales, donde la división entre evaluaciones prospectivas y retrospectivas es fundamental, como técnicas y de difusión de sus resultados. Como

apunta Novales (2023), para definir ese modelo es imprescindible establecer con precisión quién debe evaluar, cuándo, cómo y desde dónde, así como determinar quién debe ser el destinatario de los resultados de esa evaluación. Sin esa delimitación, la incorporación efectiva de la evaluación a nuestra cultura de la gestión pública es, como venimos viendo, prácticamente imposible, quedando su invocación -cada vez más habitual en la elaboración legislativa- en un mero ejercicio de retórica. Para concluir, de forma sintética esbozamos un listado de los elementos que, a nuestro juicio, debería incorporar la construcción de un marco institucional de evaluación de políticas públicas para que fuese realmente efectivo:³⁸

- Creación de una agencia u organismo independiente de evaluación de políticas públicas, adscrito preferentemente al Parlamento, con competencias sobre evaluación ex ante y ex post de programas y proyectos de inversión pública. Este organismo debe contar con independencia efectiva del Gobierno, para lo cual debe establecerse un modelo propio de financiación y selección de recursos humanos que evite cualquier dependencia del ejecutivo. La selección de personal, su formación, así como su actualización, será responsabilidad del propio organismo, preferentemente mediante procesos competitivos, abiertos a la participación de profesionales nacionales y extranjeros, de acreditada trayectoria en los ámbitos de competencia del organismo. La externalización de evaluaciones en entidades privadas o públicas debe estar motivada por razones muy tasadas y siempre condicionadas a un informe favorable del consejo rector, que se extenderá a la idoneidad de la entidad adjudicataria.
- La presidencia del organismo evaluador deberá recaer en un profesional de reconocido prestigio, con formación y experiencia acreditadas en el ámbito de la evaluación económica, dependiendo el nombramiento del Parlamento, previa presentación de una terna de candidatos propuestos por el Gobierno. El periodo de mandato deberá ser largo, y siempre superior a la duración de una legislatura -seis/ocho años-, sin posibilidad de renovación. Una vez nombrado, el presidente del organismo no podrá ser cesado sin causa justificada, que deberá validar el Parlamento. Los presidentes del organismo de evaluación estarán sometidos a limitaciones profesionales cuando concluyan su mandato, especialmente en cuanto a su incorporación a puestos de responsabilidad política.

³⁸ Los contenidos de esta propuesta están inspirados en De la Fuente (2022).

- El organismo de evaluación deberá contar con un consejo rector formado por académicos, profesionales y funcionarios de prestigio, nombrados a título personal, con trayectoria acreditada en el ámbito de actividad de la institución. En ningún caso podrán formar parte del consejo rector responsables políticos, ya sean del ámbito parlamentario o del gubernamental.
- El organismo deberá ser responsable de elaborar su programa anual de trabajo, que podrá incluir la evaluación de cualquier programa, política o proyecto de inversión pública que considere motivadamente de interés. Tras su aprobación, los planes de evaluación anuales, con respectivo calendario, deberán hacerse públicos, así como los informes sobre su grado de cumplimiento.
- En el ámbito del poder ejecutivo, debería crearse una unidad encargada del seguimiento de los resultados de las evaluaciones realizadas por el organismo de evaluación, con el fin de elaborar propuestas de reforma de la gestión pública que se puedan desprender de los informes de evaluación. Esta unidad también se encargará del impulso de tales reformas. Esta unidad podría encargar evaluaciones a la Agencia, si bien deberían estar sometidas a la aprobación de su consejo rector.
- En los departamentos gestores deben crearse unidades de control interno, cuya función será poner en marcha los mecanismos de generación de la información de calidad requerida, en todas sus facetas, por la práctica de la evaluación de políticas. Estas unidades deberán ser las encargadas de diseñar los procedimientos para la obtención de información en el diseño de las políticas públicas, así como de apoyar al organismo de evaluación en la realización de evaluaciones ex ante. La preparación de los técnicos de estas unidades, los estándares de trabajo y el funcionamiento, estará coordinado por un centro directivo con competencias horizontales en la materia (p.e. la Intervención General de la Administración del Estado), si bien la responsabilidad del buen funcionamiento de estas unidades de control interno recaerá, en última instancia, en los responsables políticos de cada departamento ministerial.
- En el caso de los procesos legislativos con elevada dotación presupuestaria, será preceptivo que las memorias de impacto normativo sean sometidas a revisión por parte del organismo de evaluación, teniendo su informe carácter suspensivo en caso de objeciones económicas motivadas.

- En el ámbito de la evaluación ex post, debería establecerse la obligación de una evaluación ex post para las leyes/proyectos con dotaciones presupuestarias por encima de un cierto umbral, con especial atención a las ayudas y subvenciones. Asimismo, se podrá fijar en el texto de una ley cuándo esta debe someterse a evaluación ex post. En ese caso, el texto de la ley deberá incluir expresamente el tipo de evaluación que ha de llevarse a cabo, además de la información necesaria para su realización, estableciéndose también los medios y procedimientos necesarios para su recogida.
- Las evaluaciones ex post a realizar deben ser incorporadas y aprobadas en los planes anuales del organismo de evaluación, quien deberá con carácter obligatorio informar anualmente sobre su realización, así como las principales conclusiones alcanzadas en cada una. El destinatario de esa información, junto con los informes elaborados, será la comisión parlamentaria que se fije al efecto. La memoria anual de actividad del organismo de evaluación, con el detalle de las actuaciones realizadas, deberá someterse a aprobación por el pleno del Congreso. Una vez aprobado, se trasladará al Gobierno para su conocimiento y adopción de medidas en relación con las recomendaciones planteadas. En caso de que el Gobierno decida no seguir tales recomendaciones, deberá dar cuenta de ello ante la comisión parlamentaria competente. Todos los informes deberán hacerse públicos, de forma íntegra, por el organismo de evaluación.
- Las administraciones públicas estarán obligadas por ley a recabar y proporcionar los datos que sean necesarios para la evaluación de las políticas públicas, así como de ponerlos a disposición, de forma operativa, del organismo de evaluación. En relación con esta información, las administraciones públicas tendrán la obligación de diseñar, elaborar y mantener ciertas bases de datos o repositorios de información de especial interés para el análisis de las políticas públicas y la investigación, cumpliendo con los estándares habituales de anonimización de la información personal. Los usuarios que soliciten el uso de estas bases de datos deberán responsabilizarse expresamente de su uso correcto para los fines declarados. El organismo de evaluación actuará como depositario de los datos y tendrá la obligación de asegurar su accesibilidad a los usuarios.
- Es necesario el establecimiento de un código de conducta del personal, así como de las entidades participantes en las actividades de evaluación, incluidas las

que pudieran ser beneficiarias de adjudicaciones de contratos de externalización, en línea con las existentes en el ámbito de la auditoría de cuentas. Este código debe garantizar la independencia de los evaluadores respecto de las entidades o instituciones responsables de las políticas, programas o proyectos evaluados.

Referencias

AFONSO, A.; JALLES, J.T.; VENÂNCIO, A. (EDS.) (2023): *Handbook on Public Sector Efficiency*, 1ª edición. Cheltenham, UK: Edward Elgar Publishing.

AFONSO, A.; SCHUKNECHT, L.; TANZI, V. (2005): "Public sector efficiency: An international comparison", *Public Choice*, 123(3-4): 321-347.

— (2010): "Income distribution determinants and public spending efficiency", *Journal of Economic Inequality*, 8(3): 367-389.

ALBI, E. (1989): *Introducción al Análisis Coste-Beneficio*, 2ª edición. Madrid: Instituto de Estudios Fiscales.

ALBI, E.; ONRUBIA, J. (2015): *Economía de la gestión pública*, Madrid: Editorial Universitaria Ramón Areces.

— (2016): "Institucionalizar la evaluación económica de políticas públicas", *Papeles de Economía Española*, 147: 50-67.

— (2018): *Supervisión, control y evaluación para una gestión pública transparente*. Madrid: Instituto de Estudios Fiscales.

ANGRIST, J.D.; IMBENS, G.W.; RUBIN, D.B. (1996): "Identification of Causal Effects Using Instrumental Variables", *Journal of the American Statistical Association*, 91(434): 444-55.

BAÑOS-ROVIRA, J. (2023): "El reto de integrar la presupuestación pública con la evaluación de políticas", *Ekonomiaz-Revista Vasca de Economía*, 103: 104-127.

BARZELAY, M. (1997): "Central audit institutions and performance auditing: A comparative analysis of organizational strategies in the OECD", *Governance*, 10 (3): 235-60.

BATTESE, G.E.; COELLI, T.J. (1992): "Frontier production functions, technical efficiency and panel data: with application to paddy farmers in India", *Journal of Productivity Analysis*, 3(1-2): 153-169.

— (1995): "A model for technical inefficiency effects in a stochastic frontier production function for panel data", *Empirical Economics*, 20 (2): 325-332.

BLUME, L.; VOIGT, S. (2011): "Does organizational design of supreme audit institutions matter? A cross-country assessment", *European Journal of Political Economy*, 27: 215-229.

BLUNDELL, R.; COSTA-DIAS, M. (2009): "Alternative approaches to evaluation in empirical microeconomics", *Journal of Human Resources*, 44(3): 565-640.

BRENT, R.J. (2006): *Applied Cost-Benefit Analysis*, 2ª edición. Cheltenham, UK: Edward Elgar.

CALIENDO, M.; KOPEINIG, S. (2008): "Some practical guidance for the implementation of propensity score matching", *Journal of Economic Surveys*, 22(1): 31-72.

CAVES, D.W.; CHRISTENSEN, L.R.; DIEWERT, W.E. (1982): "The economic theory of index numbers and the measurement of input, output, and productivity", *Econometrica*, 50 (6): 1393-1414.

COELLI, T.J.; PERELMAN, S. (1999): "A comparison of parametric and non-parametric distance functions: With application to European railways", *European Journal of Operational Research*, 117 (2): 326-339.

COMISIÓN PARA LA REFORMA DE LAS ADMINISTRACIONES PÚBLICAS (CORA) (2013): *Reforma de las Administraciones Públicas*. Ministerio de Hacienda y Administraciones Públicas y Ministerio de Presidencia. Madrid. <https://datos.gob.es/es/catalogo/e05024601-informes-cora>

DAHLBY, B. (2008): *The marginal cost of public funds: Theory and applications*. Cambridge, MA: The MIT Press.

DEBRUN, X.; HAUNER, D.; KUMAR, M.S. (2009): "Independent fiscal agencies", *Journal of Economic Surveys*, 23 (1): 44-81.

DE FRANCESCO, F.; RADAELLI, C.M.; TROEGER, V.E. (2012): "Implementing regulatory innovations in Europe: the case of impact assessment", *Journal of European Public Policy*, 19 (4): 491-511.

DE LA FUENTE (COORD.); DE RUS, G.; FERNÁNDEZ, M.; GARCÍA, M.A.; JANSEN, M.; JIMÉNEZ, S.; NOVALES, A.; ONRUBIA, J.; PÉREZ-RENOVALES, J.; SASTRE, E.; SICILIA, J. (2021): "La evaluación de políticas públicas en España: antecedentes, situación actual y propuestas para una reforma", *Fedea Policy Papers*, 2021/09. Madrid.

DE LA FUENTE, A. (2022): "Algunos comentarios sobre el proyecto de ley de institucionalización de la evaluación de políticas públicas", *Apuntes*, 2022-17, FEDEA. Madrid.

DE LA FUENTE, A.; NOVALES, A. (2023): "El marco institucional para la evaluación de políticas públicas en España", *Ekonomiaz: Revista Vasca de Economía*, 103, 36-51.

DE RUS, G. (2021): *Análisis Coste-Beneficio: Evaluación de políticas públicas y proyectos de inversión*. Barcelona: Antoni Bosch editor.

DINARDO, J. (2008): *Natural Experiments and Quasi-Natural Experiments. The New Palgrave Dictionary of Economics*. London: Palgrave Macmillan.

FINUCANE, M.M.; MARTÍNEZ, I.; CODY, S. (2018): "What Works for Whom? A Bayesian Approach to Channeling Big Data Streams for Public Program Evaluation", *American Journal of Evaluation*, 39 (1): 109-122.

FREY, B.S. (1994): "Supreme auditing institutions: A politico-economic analysis", *European Journal of Law and Economics*, 1 (3): 169-176.

FUNDACIÓN BBVA; IVIE (2023): "Stock de capital", *Esenciales*, 06/2023. Fundación BBVA e Instituto Valenciano de Investigaciones Económicas. https://www.fbbva.es/wp-content/uploads/2023/09/Esenciales06_2023_InversionPublica.pdf

GAMBAU, B.; LÓPEZ-HERRERA, C. (2023): "El papel de los Spending Review en la institucionalización de la evaluación en España: ¿qué debemos esperar?", *Ekonomiaz-Revista Vasca de Economía*, 103: 104-127.

GARCÍA-PÉREZ, J.I. (2011) (coord.): *Metodología y diseño de estudios para la evaluación de políticas públicas*, Antoni Bosch, Barcelona.

GARCÍA, R.; ONRUBIA, J.; SÁNCHEZ-FUENTES, A.J. (2019): "Is Public Sector Performance just a matter of money? The case of the Spanish regional governments", *Economics and Business Letters*, 8(2):74-84.

GERTLER, P.J.; MARTÍNEZ, S.; PREMAND, P.; RAWLINGS, L.B.; VERMEERSCH, C.M.J. (2011): *Impact Evaluation in Practice*. Washington, DC: World Bank.

GLYNN, J.J. (1985): "Value for money auditing: An international review and comparison", *Financial Accountability & Management*, 1 (2): 113-28.

GOBIERNO DE ESPAÑA (2021): Plan de Recuperación, Transformación y Resiliencia. Gobierno de España, sitio web (actualización a febrero 2024): <https://planderecuperacion.gob.es/documentos-y-enlaces>

GONZÁLEZ-PÁRAMO, J.M.; ONRUBIA, J. (2003): "Información, evaluación y competencia al servicio de una gestión eficiente de los servicios públicos", *Papeles de Economía Española*, 95: 2-23.

Haidarinejad, G.; Shekarbegi, S.; Kazemi, A.; Jamili, S. (2012): "Supreme audit court of auditors' insights on operational audit challenges", *Management Science Letters*, 2 (3): 757-762.

HARRIS, H.; HORST, S.J. (2019): "A Brief Guide to Decisions at Each Step of the Propensity Score Matching Process", *Practical Assessment, Research, and Evaluation*, 21: 1-11.

HEALD, D. (2012): "Why is transparency about public expenditure so elusive?", *International Review of Administrative Sciences*, 78 (1): 30-49.

IMBENS, G.W.; ANGRIST, J.D. (1994): "Identification and Estimation of Local Average Treatment Effects", *Econometrica*, 62(2): 467-75.

JOHNSEN, Å.; MEKLIN, P., OULASVIRTA, L.; VAKKURI, J. (2001): "Performance auditing in local government: An exploratory study of perceived efficiency of municipal value for

money auditing in Finland and Norway", *European Accounting Review*, 10 (3): 583-599.

KELLS, S.; HODGE, G. (2009): "Performance auditing in the public sector: Reconceptualising the task", *Journal of Contemporary Issues in Business and Government*, 15 (2): 36-60.

KHANDKER, S.R.; KOOLWAL, G.B.; SAMAD, H.A. (2010): *Handbook on impact evaluation: Quantitative methods and practices*; World Bank Publications, Washington, DC.

LEIBENSTEIN, H. (1966): "Allocative Efficiency vs. X-Efficiency", *American Economic Review*, 56 (3): 392-415.

MAYNE, J. (2017): "Evaluation for accountability: myth or reality?", en M.L. Bemelmans-Videc; B. Perrin y J. Lonsdale (eds.), *Making Accountability Work: Dilemmas for Evaluation and for Audit*, pp. 63-84. New Brunswick, N.J.: Transaction Publishers.

MISURACA G.; MUREDDU, F.; OSIMO, D. (2014): "Policy-Making 2.0: Unleashing the Power of Big Data for Public Governance", en M. Gascó-Hernández (ed.), *Open Government. Public Administration and Information Technology*, vol. 4. New York: Springer.

MORAL-ARCE, i. (2023): "Las potencialidades (y riesgos) de la evaluación de impacto en el aumento de la calidad de las finanzas públicas", *Ekonomiaz-Revista Vasca de Economía*, 103: 79-103.

NOVALES, A. (2023): "La evaluación de políticas públicas en España: equívocos e incumplimientos", *Anales de la Real Academia de Ciencias Morales y Políticas*, año 75, 100 (Curso 2022-2023), pp. 575-605. Madrid: Real Academia de Ciencias Morales y Políticas.

OFFICE OF MANAGEMENT AND BUDGET (2024): *Evidence and Evaluation Team*. Washington, DC: White House. <https://www.whitehouse.gov/omb/information-for-agencies/evidence-and-evaluation/>

ONRUBIA, J. (2005a): "Evaluación y gestión eficiente en el sector público: aspectos organizativos e institucionales", *Ekonomiaz-Revista Vasca de Economía*, 60: 38-57.

— (2005b): "Exigencias organizativas para una gestión pública eficiente". *Economistas*, 105: 20-28.

ONRUBIA, J.; RODADO, M.C. (2023): "La importancia de los aspectos redistributivos en la evaluación económica", *Ekonomiaz-Revista Vasca de Economía*, 103: 190-213.

ONRUBIA, J.; SÁNCHEZ-FUENTES, A.J. (2017): "How costly are public sector inefficiencies? A theoretical framework for rationalizing fiscal consolidations", *Economics e-Journal*, 11(35): 1-20.

ONRUBIA, J.; SANTÍN, D. (2017): "Gestión pública y evaluación de políticas", en R.M. Urbanos Garrido (coord.), *Lecciones para después de una crisis. Estudios en homenaje al Profesor Albi*. Madrid, Instituto de Estudios Fiscales: 329-361.

PARKER, L.D. (1986): "Value-for-Money Auditing: Conceptual, Development and Operational Issues", *Auditing Discussion Paper*, 1, Australian Accounting Research Foundation.

PEW-MACARTHUR (2013): *Results First Initiative*. The Pew Charitable Trusts, Washington, DC & The John D. and Catherine T. MacArthur Foundation, Chicago, IL. <https://www.pewtrusts.org/en/projects/archived-projects/results-first-initiative>

— (2019): *Results First Cost-Benefit Model Aids Policymakers in Funding Decisions*. The Pew Charitable Trusts, Washington, DC & The John D. and Catherine T. MacArthur Foundation, Chicago, IL. https://www.pewtrusts.org/-/media/assets/2019/05/resultsfirst_overview_of_the_results_first_cost_benefit_model_v1.pdf

PREST, A.R.; TURVEY, R. (1966): "Cost-Benefit Analysis: A Survey", *Surveys of Economic Theory*, vol. III. London: Palgrave Macmillan.

RODRÍGUEZ-VIVES, M. (2019): "The quality of public finances: where do we stand?", *Economics and Business Letters*, 8(2): 97-105.

SANTÍN, D. (2015): "Evaluación de las políticas públicas", en E. Albi y J. Onrubia, *Economía de la Gestión Pública. Cuestiones Fundamentales*, Cap. 4, pp. 113-152, Editorial Universitaria Ramón Areces, Madrid.

THÖNE, M. (2023): "The quality of public finances", en A. Afonso, J.T. Jalles y A. Venâncio (eds.), *Handbook on Public Sector Efficiency*. Cheltenham, UK: Edward Elgar Publishing, pp. 32–43.

UK TREASURY (2020): *Magenta Book*. London: UK Treasury, UK Government. Website (10/2/2024): <https://www.gov.uk/government/publications/the-magenta-book>

VALIÑO, A. (2023): "La evaluación económica en el análisis de las políticas públicas: el estado del arte en España", *Ekonomiaz-Revista Vasca de Economía*, 103: 159-189.

WACHS, D.; ROMAN-URRESTARAZU, A.; BRAYNE, C.; ONRUBIA, J. (2020): "Dependency ratios in healthy ageing", *BMJ Global Health*, 5(4), e002117.

WILLIAMS, A. (1972): "Cost-benefit analysis: bastard science? And/or insidious poison in the body politick?", *Journal of Public Economics*, 1: 199–225.

WILLIAMS, A.; EVANS, R.; DRUMMOND, M. (1987): "Quality-adjusted life-years", *The Lancet*, 329(8546): 1372-1373.