

LA EVALUACIÓN DE POLÍTICAS PÚBLICAS EN ESPAÑA: EQUÍVOCOS E INCUMPLIMIENTOS*

Por el Académico de Número
Excmo. Sr. D. Alfonso Novales Cinca**

Evaluar una política pública consiste en comparar el bienestar social con y sin la aplicación de la política su bienestar social con su coste social; como criterio general una política pública debería aplicarse si su contribución al bienestar social supera a su coste social [Rus (2023)]. Para concretar esta idea se han propuesto muchas definiciones del concepto de evaluación; la OCDE concibe la evaluación de políticas públicas como «la valoración estructurada y basada en evidencia observable, del diseño, la implementación o los impactos de una intervención pública planificada, en curso o completada» [OECD (2020)]¹.

1. ¿POR QUÉ ES TAN RELEVANTE LA EVALUACIÓN DE POLÍTICAS PÚBLICAS?

Las políticas públicas utilizan recursos públicos, ya sean aportados por los contribuyentes actuales mediante impuestos, o generados por endeudamiento, al que habrán de hacer frente los contribuyentes futuros. Los recursos públicos tienen un coste explícito, y es peligrosamente sencillo justificar su

* Esta ponencia tiene su origen en los trabajos: DE LA FUENTE *et al.* (2021), y DE LA FUENTE y NOVALES (2023).

** Sesión del día 30 de mayo de 2023.

¹ «La evaluación de políticas públicas es un proceso sistemático y razonado de generación de conocimiento, a partir de la recopilación, análisis e interpretación de información, encaminado a la comprensión global de una intervención pública –sea esta una política, plan, programa o norma–, para alcanzar un juicio valorativo, basado en evidencias, respecto de su diseño, puesta en práctica y efectos (resultados e impactos).» (IEPP, 2021).

incremento, porque siempre pueden encontrarse necesidades alternativas que atender. Hay que recabarlos de algún modo y, en este proceso, los gobiernos no quieren generar grandes tensiones, ni cargar de compromisos económicos a la sociedad y, además, deben tener en cuenta aspectos de equidad intergeneracional.

Puesto que son costosos y compiten entre usos alternativos², la preocupación por un uso eficiente de los recursos públicos ha estado siempre presente en la legislación. El artículo 31.2 de nuestra Constitución requiere la exigencia de que la programación y ejecución del gasto público respondan a criterios de eficiencia y economía. La Ley 40/2015 de Régimen Jurídico del Sector Público (art. 3) exige el control de la gestión y la evaluación de resultados de políticas públicas, a través de las inspecciones de servicios y la IGAE.

Por tanto, dos razones justifican la necesidad de evaluar los resultados de las políticas públicas: en primer lugar, la información proporcionada por los procesos de evaluación es imprescindible para mejorar la eficiencia y la eficacia en el uso de los recursos públicos, tal como pide la Constitución. Una segunda motivación es la rendición de cuentas a los ciudadanos, quienes proveen los recursos y para cuyo bienestar se han escogido, supuestamente, las políticas implementadas.

Los análisis comparados muestran que la evaluación rigurosa de políticas públicas es todavía una asignatura pendiente en España³, lo cual no hace, a mi juicio, sino reflejar cierta despreocupación por una rigurosa rendición de cuentas de los gestores públicos. Esto no es bueno; no conviene ignorar que la transparencia, la evaluación de las políticas públicas y la rendición de cuentas de los gestores públicos, constituyen tres pilares sobre los que se puede juzgar la calidad de una democracia. En los tres estamos haciendo ciertos progresos en España pero, en algunos aspectos relevantes, considero que los avances son más aparentes que reales⁴.

² Una de las definiciones convencionales de Economía más extendidas es la de L. Robbins (1932), quien se refiere a la economía como la «Ciencia que estudia la utilización óptima de los recursos escasos, susceptibles de usos alternativos».

³ Stockmann, Meyer y Taube (2020) asignan a España un nivel medio-bajo en la evaluación de políticas públicas desde una perspectiva institucional en términos comparados. Albi y Onrubia (2016, p. 59) opinan: «Cuatro décadas después de la inclusión en la LGP de 1977, la LGP 2003 y de su reconocimiento constitucional como principios orientadores del gasto público (art. 31.2 de la Constitución Española de 1978), el desarrollo institucional, regular y sistemático de su evaluación sigue siendo marginal».

⁴ Los resultados obtenidos en la evaluación de una política deben someterse a una rendición de cuentas por parte de los responsables de los departamentos promotores de dicha política. Por primera vez tenemos en esta legislatura un sistema regular de rendición de cuentas, con información clara acerca del cumplimiento de compromisos (<https://www.lamoncloa.gob.es/Paginas/cumpliendo/index.aspx>). Pero la mayoría, si no la totalidad, de dichos compromisos es de naturaleza cualitativa, y no es claro que el método seguido para calificar su grado de cumplimiento sea el más deseable, si bien está avalado por un grupo de reconocidos académicos.

2. ¿QUÉ TIPOS DE EVALUACIÓN PODEMOS DISTINGUIR Y CUÁLES SON SUS OBJETIVOS?

Hay varios tipos de evaluación porque, en distintos momentos de la aplicación de una política, la evaluación puede cumplir con diferentes objetivos. Inicialmente, la evaluación *ex-ante* debe justificar rigurosamente la elección de una determinada intervención pública. Por tanto, debe considerar políticas o actuaciones alternativas que pudiesen haber conducido al cumplimiento del mismo objetivo que la política escogida. En el proceso de evaluación *ex-ante* debe consultarse el diseño de una política con expertos: educadores, juristas, investigadores, instituciones (como ya sucede con los actuales informes preceptivos o consultivos del Consejo de Estado, y conviene analizar las experiencias similares en otros países en términos de la política que se propone. De acuerdo con la Ley de Transparencia, en el caso de una política respaldada por una norma, la información recibida en estas consultas debe formar parte de su Memoria de Análisis de Impacto Normativo, a las que luego me referiré.

Un segundo tipo de evaluación consiste en hacer un seguimiento de la aplicación de una política y de los resultados que se van alcanzando en dicho proceso. Para ello, es conveniente establecer previamente un escenario temporal acerca de su puesta en práctica, así como de sus potenciales resultados, incorporando en dicho escenario la incertidumbre de manera adecuada. Seguir el desarrollo y los primeros resultados de una política permite introducir correcciones, tanto si se deben a unos resultados parciales muy distintos de lo esperado, como si la política no puede desarrollarse al ritmo previsto. Permitirá también analizar las causas de las posibles desviaciones, para actuar sobre ellas o para modificar los objetivos inicialmente trazados. Aporta una flexibilidad que permite mejorar la eficacia de las políticas públicas, y proporciona asimismo información de la mayor utilidad para el diseño de políticas futuras.

Por último, la evaluación *ex-post* o evaluación de impacto analiza los efectos de una política, pero no debe confundirse con una mera contabilidad de lo realizado. Constatar los kilómetros de carretera o vía férrea construidos como fruto de una política de reactivación de determinado territorio, y relacionarlos con el presupuesto ejecutado, no constituyen una evaluación de dicha política. El coste por kilómetro es una evaluación de la eficacia en costes de la construcción de la infraestructura viaria, pero no una evaluación de la política. El modo en que se diseñó la evaluación del Plan Anual de Política de Empleo dentro de la Estrategia de Apoyo Activo al Empleo 2021-2024 es un claro ejemplo de este tipo de confusión [Cebrián *et al.* (2021)]⁵.

⁵ Tras aplicar una política de empleabilidad, conocer cuántas personas se han formado o a cuántas se les ha ofrecido un itinerario de inserción es relevante; también lo es saber en qué medida las personas registradas en los servicios públicos de empleo transitan a la ocupación. Pero estos resultados no implican que la política haya tenido impacto. Tras haber pasado por una acción formativa o por un itinerario se puede

Cuanto mejor sea la evaluación ex-ante, mejor serán la evaluación ex-post. Los tres tipos de evaluación deben verse como complementarios, aun siendo de diferente naturaleza. Además de ser una obligación de responsabilidad hacia la sociedad, la evaluación es, en sus tres formas complementarias, una herramienta crucial para la mejora continua de las políticas públicas, tanto en términos de su conceptualización y diseño, como de la eficacia y eficiencia en su gestión y ejecución, y de los métodos de análisis de sus resultados e impactos.

Para evaluar, una política no debe limitarse a establecer objetivos cualitativos como pudiera ser «mejorar la calidad de la administración», «mejorar los resultados escolares», o «reducir el paro». La evaluación queda devaluada cuando una política carece de objetivos cuantificables, como una determinada reducción de la tasa de abandono escolar, o un determinado incremento del empleo de mayores de 50 años. La ausencia de objetivos cuantitativos es habitual en el diseño de políticas públicas en España, como se ha visto en los muchos programas de actuación, Planes y Estrategias lanzados en los últimos años. Puede deberse a la resistencia de los promotores de la actuación a someterse a un posible incumplimiento de objetivos, pero esta reticencia es algo que debe terminar. La idea es que haya una rendición de cuentas donde se confronten los resultados obtenidos con los esperados, y se puedan explicar las posibles desviaciones.

3. ¿QUIÉN DEBE EVALUAR? ¿DONDE DEBE UBICARSE LA EVALUACIÓN?

Los departamentos gestores de una política tienen mejor información sobre su diseño e implementación, así como sobre las necesidades que trata de resolver, y conviene que realicen una evaluación de tipo *interno*. Pero, generalmente, no cuentan con el personal especializado para realizar evaluaciones sofisticadas y, evidentemente, se enfrentan a problemas de incentivos para evaluar sus propias iniciativas y actuaciones.

En segundo lugar, conviene someter las políticas públicas a una evaluación *externa* a cargo de agencias o equipos especializados, que pueden ser más rigurosas e imparciales, según cuál sea su ubicación concreta en la administración. Cuanta más autonomía tenga la agencia evaluadora respecto del

haber obtenido un empleo, pero ello no implica que se deba a que se ha participado en dichas políticas. Las cifras que se obtengan pueden estar más determinadas por la estructura económica de la zona en que se llevan a cabo las políticas, y por su coyuntura, que por la eficacia de las políticas. De modo similar, medir el grado de empleabilidad o el desempeño laboral no es tarea fácil: hay que poder determinar qué se entiende y cómo se mide, porque no solo depende de que se logre salir del desempleo, sino del tiempo que se tarda en lograrlo, de la proporción de tiempo que el individuo está empleado, el nivel de ganancias acumuladas, el tiempo transcurrido hasta el primer empleo, del tipo de contrato con el que se accede, de la duración de dicho contrato, del nivel de cualificación y su adecuación con el empleo, entre otras [Cebrián *et al.* (2021)].

poder ejecutivo, mayores serán los incentivos para llevar a cabo una evaluación rigurosa de una determinada actuación. Por otro lado, cuanto más proximidad tenga con el ejecutivo, mayor será la influencia que pueda ejercer para que su evaluación conduzca a mejoras en el diseño de futuras políticas, si bien esto requiere un firme compromiso político con la evaluación y con la utilización de sus resultados. En todo caso, las distintas unidades de evaluación que puedan existir deben verse como complementarias, más que sustitutivas, y deben tener claramente delimitadas sus competencias y responsabilidades.

Por último, es importante la actividad de evaluación que desarrollen académicos e investigadores, tanto por su mayor disponibilidad para estar al corriente de las metodologías más novedosas y eficaces, como por ofrecer una valoración independiente que pueda confrontarse con las obtenidas desde otras instituciones. El debate a que esta competencia pueda dar lugar, no solo entre nuestros representantes políticos, sino también entre expertos e investigadores, es la mejor manera de garantizar que las políticas se diseñen y ejecuten cada vez con mayor rigor.

4. METODOLOGÍAS DE EVALUACIÓN DE POLÍTICAS PÚBLICAS⁶

Evaluar una política consiste en comparar el nivel de bienestar social con y sin la actuación pública; lamentablemente, este es un deseo irrealizable, pues el bienestar social no es observable. Además, por el Teorema de Imposibilidad de K. Arrow, tampoco podemos agregar preferencias individuales, si pudiésemos averiguarlas, en un indicador de bienestar social. Durante décadas, la metodología más extendida, especialmente en la evaluación de infraestructuras, ha sido la que se conoce como análisis Coste-Beneficio, que ha tratado de relacionar el beneficio social generado por la actuación pública, con su coste. Este análisis se enfrenta a tres dificultades: los beneficios sociales son en gran medida no observables; como se dejan notar a lo largo del tiempo, hay que descontarlos adecuadamente, y pequeñas variaciones en la tasa de descuento social utilizada pueden alterar drásticamente los resultados; además, con antelación a la aplicación de la política, dichos beneficios son inciertos. Estas dificultades, que son en buena medida inherentes al análisis económico, requieren un cuidadoso tratamiento.

Otro grupo de metodologías trata de medir la eficiencia productiva, que consiste en evaluar si, con las cantidades de inputs utilizadas en un proceso productivo podrían obtenerse más unidades del output, o si para obtener las unidades de output observadas, podrían utilizarse menos unidades de los inputs

⁶ Una versión completa de esta sección aparece en el Apéndice 1.

utilizados en el proceso. Este enfoque es útil para comparar la eficiencia productiva entre unidades homogéneas que ofrecen un mismo servicio público, permitiendo establecer rankings entre ellas, e identificando las causas de su posible ineficiencia.

Para la evaluación de impacto, que tiene carácter ex-post, ha cobrado gran relevancia en los últimos años en el análisis económico la utilización de experimentos, de los que nos habló nuestra compañera Mari Paz Espinosa en su discurso de ingreso en esta Real Academia. Este enfoque permite, bajo determinadas condiciones, estimar el impacto causal de una política mediante la comparación de resultados en unidades (que pueden ser personas, regiones) en las que se ha aplicado la política, con otras similares en las que no se ha aplicado. Este es un área donde se han producido recientemente considerables avances que han motivado, en parte, la concesión del Premio Nobel en Economía, los años 2019 a Banerjee, E. Duflo y 2021 a D. Card, J. Angrist, y G. Imbens.

5. ¿QUÉ RECORRIDO HA EFECTUADO LA EVALUACIÓN DE POLÍTICAS PÚBLICAS EN ESPAÑA?

La evaluación comenzó a ser tomada en consideración en España tras el ingreso en la actual UE. En 2007 se creó la Agencia Estatal de Evaluación de la Calidad de los Servicios y de las Políticas Públicas (AEVAL). Contó con escaso personal cualificado y era excesivamente dependiente del poder ejecutivo: estaba adscrita al Ministerio de Administraciones Públicas; su presidente y sus consejeros eran nombrados por los Ministerios, siendo en su mayoría altos cargos, y su plan de evaluación era aprobado por el Consejo de Ministros. Realizó estudios sobre la «calidad de los servicios públicos», pero apenas hizo nada en términos de evaluación de impacto de programas y políticas públicas [Albi y Onrubia (2018), p. 251].

Su supresión en 2017, nunca adecuadamente justificada, frustró las expectativas de avance en este ámbito. Fue sustituida por un Instituto para la Evaluación de Políticas Públicas, con rango de subdirección general, dentro de la Secretaría de Estado de Función Pública. Desde entonces, este Instituto ha desarrollado una actividad limitada, sin abundar las evaluaciones rigurosas de impacto o los análisis Coste-Beneficio de programas de cierta entidad.

En octubre de 2012 se creó la Comisión para la Reforma de la Administración (CORA)⁷. Su informe de 2013 introdujo el requisito de que toda nueva

⁷ «Las 217 medidas adoptadas o propuestas en su informe de 2013 carecen del soporte analítico suficiente, ni en la mayoría de los casos son fruto de estudios técnicos rigurosos basados en métodos científicos».

normativa viniera acompañada de una Memoria de Análisis de Impacto Normativo (MAIN), ya introducidas en la Ley del Gobierno (Ley 50/1997), y reguladas por el RD 1083/2009 y el DL 931/2017⁸. Dicha Memoria está orientada hacia una evaluación *ex-ante*, pues debe incluir una justificación de la nueva norma, analizando su impacto jurídico, e indicando normas que puedan quedar derogadas. También debe incorporar el impacto económico y presupuestario de la nueva norma, describir los informes consultados para su diseño y redacción, las contribuciones recibidas de los expertos e instituciones consultadas y los resultados de los procesos de opinión pública y de información pública a los que debe someterse un anteproyecto de ley, antes de su aprobación de acuerdo con la Ley del Gobierno (art. 26).

Esa misma Ley del Gobierno establece la aprobación de un Plan *Normativo* Anual detallando las iniciativas legislativas que habrán de ser elevadas cada año al Consejo de Ministros para su aprobación, e indicando cuáles de ellas habrán de someterse a una evaluación *ex post*⁹. De modo complementario, el *Informe Anual de Evaluación*, que debe elaborar el Ministerio de la Presidencia, refleja el grado de cumplimiento del Plan Anual Normativo del año anterior, así como las conclusiones del análisis de las normas que hayan sido evaluadas, atendiendo a su eficacia, eficiencia y sostenibilidad, pudiendo dar lugar a recomendaciones para la modificación o derogación de las normas evaluadas. Sin embargo, a pesar de su nombre, ni el Plan Normativo ni el Informe de Evaluación han llegado a presentar realmente evaluaciones. Son, simplemente, un índice de normativas a someter a aprobación, y un chequeo de sí, efectivamente, se han propuesto y han sido aprobadas. Tanto el Plan Normativo como el informe de Evaluación deben aprobarse por el Consejo de Ministros y publicarse en la página web de Transparencia antes del 30 de abril, lo cual no se ha producido este año, y también se ha incumplido en años anteriores¹⁰.

ficos o profesionales homologables. Su último informe trimestral data del segundo trimestre de 2016 y tampoco se han vuelto a producir informes anuales. Como ocurrió con el Plan de Modernización de la Administración del Estado de 1992, los avances registrados en los informes emitidos recogen algunas actuaciones de escasísima relevancia, impropias de aparecer incluso en una estrategia rigurosa de reforma o de modernización» [Sebastián (2019), cap. 4 y Apéndice].

⁸ En el caso de los decretos y real decreto-leyes, la ley permite que la memoria técnica tenga carácter abreviado, pero exige que incluya, entre otras cosas, el análisis del impacto presupuestario del texto legislativo al que acompaña (DL 931/2017, art. 3).

⁹ En la reformulación establecida en la disposición final tercera de la Ley de Régimen Jurídico del Sector Público (Ley 40/2015, art. 25). El Plan Anual Normativo y el Informe Anual de Evaluación están regulados por el Real Decreto 286/2017.

¹⁰ En el momento de cerrar esta ponencia (14 de mayo de 2023), ni el Plan Normativo anual para 2024 ni el Informe Anual de Evaluación Normativa correspondiente a 2022 han sido publicados en la página web de Transparencia, como requiere que se haga antes del 30 de abril de cada año la Ley de Gobierno (Ley 50/1997, art. 25 y 28). El Informe Anual de Evaluación Normativa de 2022 ya fue aprobado en el Consejo de Ministros de 25 de abril de este año, pero no el Plan Normativo anual para 2024. Este incumplimiento de fechas no es nuevo: el Plan Anual Normativo para 2023 fue aprobado el 31 de enero de 2023, nueve meses después de lo requerido. En el Informe de Evaluación Normativa de 2021 se decía (p. 10): «Por ello, merece destacarse que esta es la segunda vez, desde que entraron en vigor en 2016 las previsiones normativas que prescriben la aprobación de un Plan Anual Normativo y de un Informe Anual de Evaluación

Como puede verse, la responsabilidad sobre la evaluación, muy desvirtuada conceptualmente, ha estado siempre en manos del ejecutivo. De hecho, generalmente la evaluación se asigna a los propios departamentos responsables del diseño de las normas o de su gestión, lo que plantea problemas evidentes de incentivos. Si a esto añadimos la falta de personal cualificado para realizar evaluaciones rigurosas de políticas complejas, y la escasa disponibilidad de la información precisa para llevarlas a cabo, no resulta sorprendente que las memorias de impacto normativo sean frecuentemente meros trámites formales¹¹ que se preparan con el núcleo del proyecto de ley ya prácticamente cerrado, aun a pesar de la revisión que de las mismas hace la Oficina de Coordinación y Calidad Normativa¹². Por tanto, rara vez sirven como filtros para descartar proyectos de baja calidad o para mejorar su diseño.

En ocasiones, se han asumido compromisos en materia de evaluación que no se han cumplido: dentro de la Estrategia Española de Activación para el Empleo 2017-2020, la Conferencia de Presidentes del 17 de enero de 2017 acordó establecer evaluaciones, anuales o bienales, del cumplimiento de objetivos y del desempeño de las políticas activas de empleo, así como de la calidad, impacto, eficacia y eficiencia del conjunto del sistema de formación profesional para el empleo, a las que se daría difusión pública. También se propuso implantar una metodología común para la evaluación de los programas y medidas concretos recogidos en el Plan Anual de Política de Empleo, pero se ha avanzado poco en el cumplimiento de tales compromisos [Cebrián *et al.* (2022)]. Sin embargo, en la Estrategia Española de Apoyo Activo al Empleo 2021-2024 (BOE A-2021-20185) y en la resolución de aprobación del Plan Nacional de Empleo para 2022 (BOE 4/5/2022) se mencionan que en su redacción se ha tenido en cuenta la primera y segunda evaluación realizada por la Red de Servicios Públicos Europeos. Lamentablemente, dichas evaluaciones no parecen estar disponibles. Tampoco están accesibles las múltiples evaluaciones y autoevaluaciones, mencionadas en el primero de estos documentos, de la propia Estrategia de apoyo activo al empleo y de los diversos programas que en ella se consideran. Es fácil acceder, por el contrario, a la evaluación que la AIReF hizo en el Spending Review (Revisión del gasto) de 2018 de las políticas activas de empleo en España.

Normativa, que el Gobierno aprueba dicho Informe y procede a su publicación en el Portal de la Transparencia para conocimiento de los ciudadanos; siendo la primera vez que dicha aprobación se realiza en el plazo establecido en su normativa reguladora. Con ello, se sigue avanzando en el propósito de dotar de la máxima transparencia y accesibilidad a los resultados de la acción de gobierno en el ámbito de la elaboración y aprobación de normas».

¹¹ Véase, por ejemplo, DE LA FUENTE (2023).

¹² De acuerdo con el Real Decreto 1081/2017 (art. 2). La Oficina de Coordinación y Calidad Normativa no mantiene una página web abierta al público, ni se conocen los informes que realiza sobre las Memorias de Análisis de Impacto Normativo de los distintos anteproyectos de ley.

Diversas instituciones realizan tareas de evaluación en España. Entre ellas, el Tribunal de Cuentas, la IGAE, Oficina Nacional de Auditoría (dependiente de IGAE, a diferencia de otros países), el Instituto de Evaluación de Políticas Públicas, y la Oficina Nacional de Evaluación, creada en 2021 para supervisar los contratos públicos¹³. Esta relación de instituciones ilustra la confusión existente en nuestra normativa entre tareas de evaluación de impacto y tareas de auditoría presupuestaria¹⁴.

El intento más reciente de reforzar la evaluación de las políticas públicas han sido los llamados Spending Reviews (revisiones de gasto) realizados por la Autoridad Independiente de Responsabilidad Fiscal, AIReF, con el objetivo de mejorar la eficiencia del gasto público. Estas revisiones se realizan por encargo del Gobierno, motivado a su vez por una recomendación de la Comisión Europea que surgió en torno a la discusión del Plan Presupuestario para 2017¹⁵. En uno de los ejemplos más llamativos, utilizando la última información disponible, correspondiente a finales de 2021, AIReF (2022) estimaba que el Ingreso Mínimo Vital solo había llegado al 40% de sus beneficiarios potenciales, que debían ser unos 700.000 hogares, sin incluir el País Vasco y Navarra, y solo se había ejecutado un 56% de su presupuesto (BdE, *Informe anual 2023*, p. 132).

También las CC.AA. se van sensibilizando hacia la evaluación de sus políticas públicas. Varias han aprobado una normativa para desarrollar un sistema público de evaluación, o incluyen referencias a la evaluación en sus normas

¹³ Esta última está encargada de analizar la sostenibilidad financiera de los contratos de concesiones de obras y de servicios públicos, incluida la valoración de la rentabilidad de cada concesión.

¹⁴ «A pesar del amplio espectro legal, la realidad muestra que estamos ante una concepción de la auditoría eminentemente instrumental y escasamente institucional, con una consideración residual en el sistema de control de la gestión pública». «Si bien los controles de legalidad y de regularidad contable, ya sea a través de la práctica de la función interventora de fiscalización previa o del ejercicio del denominado control financiero permanente han sido ejercidos por el Tribunal de Cuentas o la IGAE, [...] la realización de controles de eficiencia, economía y eficacia sigue siendo una asignatura pendiente. La consolidación presupuestaria y el saneamiento de las finanzas públicas priman por encima de la persecución de la mejora de la eficiencia y la eficacia de los programas públicos» [Albi y Onrubia (2016), pp. 58, 59 y 64].

¹⁵ La independencia de al AIReF se garantiza, entre otros mecanismos, por su forma de financiación mediante una tasa y el modo de selección de su presidenta, que requiere la mayoría absoluta del Parlamento y se realiza fuera de las legislaturas. De las tres fases inicialmente previstas, hasta el momento se han completado dos, la última en octubre de 2020, en las que se han incluido evaluaciones de diversas políticas de gasto estatales y autonómicas, algunas con un elevado peso presupuestario, siguiendo procedimientos mucho más cuidadosos de lo que había sido habitual hasta el momento en nuestro país. El Spending Review actualmente en curso se completará con una tercera fase de revisión del gasto del conjunto de las AAPP, centrada en los préstamos y ayudas públicas a las empresas y la gestión de los residuos urbanos, que se preveía finalizar antes de julio de 2022 pero cuyos resultados todavía no se han presentado en el momento de escribir estas líneas. Los planes evaluados hasta ahora incluyen las políticas activas e incentivos al empleo, algunos componentes del gasto sanitario, las becas y ayudas universitarias y para la contratación de personal investigador, los beneficios fiscales, la estrategia y procedimiento en la concesión de subvenciones, y ciertas inversiones ferroviarias. Los estudios pertinentes se llevaron a cabo por personal propio de la AIReF en colaboración con grupos externos de consultores y académicos y con diversas administraciones públicas, dada la escasez de personal propio. En la AIReF el personal dedicado a la evaluación no llega a una decena, y el Instituto para la Evaluación de Políticas Públicas cuenta con once personas a tal efecto (Casado y Del Pino, 2022).

de transparencia, participación o gobierno abierto¹⁶, y al menos ocho CC.AA. han hecho encargos de evaluación de algunas de sus políticas a la AIREF.

Conviene valorar también que en universidades y centros de investigación España cuenta con un buen número de expertos en evaluación, que han hecho aportaciones muy relevantes tanto en el desarrollo y aplicación de metodologías de evaluación, como en su aplicación a casos concretos, especialmente en los ámbitos educativo, tributario, y de políticas de empleo [véanse VV.AA. (2021) y VV.AA. (2022)]¹⁷.

En resumen, puede decirse que los males tradicionales de nuestra práctica de evaluación de políticas públicas son: 1) la escasez de medios humanos, en número, cualificación, y experiencia, 2) la excesiva dependencia del ejecutivo, generando de este modo unos incentivos erróneos, 3) insuficiente presupuesto, 4) una concepción mecánica de la evaluación, bajo un concepto distorsionado, que la identifica en ocasiones con el control contable, financiero, de regularidad/legalidad, 5) arbitrariedad en los mecanismos de selección de políticas a evaluar, 6) la escasa utilización de los resultados de las pocas evaluaciones que se hacen, 7) la ausencia de una vocación evaluadora por parte de los distintos gobiernos.

En este contexto se aprueba la Ley de Institucionalización de la Evaluación de Políticas Públicas, Ley 27/2022, de 20 de diciembre de 2022, a la que me referiré en lo sucesivo como «Ley de Evaluación».

6. ¿CUÁLES SON LAS PRINCIPALES CARACTERÍSTICAS QUE INCORPORA LA NUEVA LEY?

El preámbulo de la Ley describe que más de un centenar de unidades públicas de diferente naturaleza tienen funciones o competencias relacionadas con la evaluación, pero menos de una cuarta parte evalúan las políticas públi-

¹⁶ Los Órganos de Control Externo de las Comunidades Autónomas, cuando se han creado por el Estatuto autonómico u otra legislación posterior, tienen atribuciones sobre el control externo de las cuentas y de la gestión económico-financiera de las instituciones y entidades del Sector público autonómico y local, siendo compatibles con los análisis del Tribunal de Cuentas. La disponibilidad de unidades de evaluación es muy heterogénea entre distintas CC.AA., las cuales se enfrentan a las mismas limitaciones que la administración central, necesitando reforzar la formación especializada en empleados públicos, así como los sistemas de información y de utilización de datos.

¹⁷ Entre ellas, el Ingreso Mínimo Vital, y el período de derecho de las prestaciones de los trabajadores mayores en el año 2012, el contrato en prácticas profesionales, los programas Emple@Joven y Emple@30+, desarrollados en Andalucía en 2016-18, o el programa de Ayuda a la Dislexia de la Comunidad de Madrid dirigido a alumnos de primaria en el curso 2018-2019. Por otra parte, la propuesta de evaluar el desempeño de los empleados públicos suscitó fuertes reticencias hasta el punto de ser eliminada del anteproyecto de ley de Función Pública.

cas contenidas en estrategias, planes o programas. La ley se justifica para tratar de resolver «la dispersión y falta de uniformidad de la práctica de la evaluación, disparidad de criterios a la hora de definir la evaluación de políticas públicas y sus contenidos, ya que se incluye en dicho concepto una gran variedad de actividades de seguimiento, auditoría, supervisión y análisis» (Preámbulo II).

Esta situación se afronta en la ley creando un entramado de organismos e informes. Se crean 3 organismos estatales: la Agencia Estatal de Evaluación de Políticas Públicas, la Comisión Superior de Evaluación, que debe coordinar los distintos ministerios entre sí y con la AGE, y el Consejo General de Evaluación, de carácter consultivo y asesor, en el que estará representada la sociedad civil. Además, se crearán unidades de coordinación de la evaluación de políticas públicas en cada Ministerio. En cuanto a qué evaluar, la Ley establece un Plan de Evaluaciones Estratégicas (PEE) del Gobierno, de carácter cuatrienal, que afectará a las políticas públicas consideradas relevantes, así como Planes departamentales de Evaluación, uno por ministerio, de carácter bienal, que se presentarán a la Comisión Superior de Evaluación, si bien la Ley no especifica con qué objeto se presentarán a dicha Comisión. Se introduce asimismo un «Sistema de Indicadores Comunes» para la evaluación.

La Agencia Estatal de Evaluación de Políticas Públicas está llamada a jugar un papel central, pues tiene asignadas múltiples y complejas tareas: 1) elabora la propuesta del PEE del Gobierno, 2) realiza el análisis de Evaluabilidad de las políticas públicas incluidas en el PEE, al que luego me referiré, 3) aprueba los Planes departamentales de evaluación, 4) elabora la Memoria anual del estado de la evaluación de políticas públicas, a presentar a las Cortes Generales, 5) verifica el cumplimiento de recomendaciones de las evaluaciones, 6) formula metodologías, 7) diseña, gestiona y hace un seguimiento del Sistema de Indicadores, además de diversas responsabilidades en formación, coordinación, cultura, representación, portal temático, relaciones con Universidades, entre otras.

Sorprende que, a pesar de la relevancia de la Agencia Estatal, la ley no describe la composición de sus órganos de dirección, las características de su personal, ni el proceso de selección del mismo. La ley le otorga un «alto grado de independencia y autonomía», a pesar de ubicarla en la Secretaría de Estado de la Función Pública. Tampoco parece buena idea para garantizar su independencia que el Plan de Evaluaciones Estratégicas, que debe elaborar esta agencia esté sujeto a posible modificación y aprobación por parte del Consejo de Ministros.

La ley genera también una notable carga administrativa, sin dejar clara su utilidad. Requiere realizar un análisis de Evaluabilidad «de (todos) los anteproyectos, programas, planes estratégicos y toda iniciativa que desarrolle políticas públicas» (11.1), a realizar por la Agencia Estatal de Evaluación, «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 accesibi-

lidad del sistema de información y evaluación previsto». Parece ésta una tarea inasumible, tanto por el número de informes a realizar, como por su complejidad, que está descrita minuciosamente en la ley. Los planes aprobados podrán tener preferencia en su presupuestación, pero no se especifica que sucederá si un programa, plan, o propuesta de ley no supera el análisis de Evaluabilidad¹⁸. Todo sugiere que puede convertirse en un trámite burocrático más, superado de modo mecánico.

En segundo lugar, cada política analizada tendrá un Informe Final de Evaluación que «recogerá, al menos, el enfoque y la metodología utilizado en la evaluación, los resultados del análisis, y las conclusiones y recomendaciones. El equipo evaluador remitirá un borrador al órgano responsable de la política pública para que realice las observaciones oportunas que el equipo evaluador valorará y, en su caso, incorporará en el Informe Final de Evaluación» (art. 22). Las recomendaciones de este informe, junto con las medidas adoptadas para solventar las deficiencias apuntadas, deben incluirse dos meses después en un último Informe de Resultados¹⁹.

En cuanto al contenido conceptual de la ley, su comienzo es prometededor, pues el preámbulo motiva correctamente la evaluación de políticas: «La evaluación favorece la toma de decisiones informada, propone eventuales correcciones, contribuye a la eficiencia y eficacia de la acción pública, y es un elemento de profundización de la democracia, al servir para la rendición de cuentas ante la ciudadanía». Pero su lectura detallada va debilitando las expectativas iniciales, por tres razones principales.

¹⁸ (Art. 11) «Con el fin de verificar la calidad y precisión del diagnóstico realizado, la viabilidad, coherencia y consistencia del diseño elaborado, la claridad, suficiencia y accesibilidad del sistema de información y evaluación previsto, así como si existen suficientes capacidades, recursos y compromisos para que la evaluación pueda implementarse. Se valorarán, como mínimo, los siguientes elementos: La calidad de la planificación, la existencia de un sistema de información, seguimiento, y evaluación que posibilite la monitorización periódica y sistemática de la ejecución de la política pública, y la gobernanza de la misma. El análisis de evaluabilidad podrá realizarse por un equipo evaluador interno o externo al órgano responsable de la política pública objeto de evaluación, y se documentará en un Informe de Evaluabilidad que contendrá qué elementos de la política pública necesitan mejorarse y sus posibles alternativas. Las recomendaciones del Informe de Evaluabilidad serán tenidas en cuenta por el órgano responsable de la política pública, debiendo explicar, en su caso, su separación de las mismas, y será remitido a la Unidad de Coordinación de Evaluación de políticas públicas en el departamento ministerial para su seguimiento».

¹⁹ Se enfatiza en la ley la creación de un Sistema de Indicadores Comunes, para su utilización en los procesos de evaluación, pero no se explica qué se entiende por un «indicador común». Anuncia también la constitución de un «Grupo de Trabajo de indicadores para colaborar en la definición y construcción de los indicadores generales de evaluación ex ante contenidos» (art. 28). Estas referencias son difíciles de entender, pues distintas políticas pueden requerir indicadores específicos, no siendo sencillo establecer el sistema de indicadores comunes al que se hace referencia en la ley. Además, la mera referencia un sistema de indicadores comunes, tiene resonancias de una concepción mecánica de la evaluación, más centrada en la eficiencia de costes o en la medición de kilómetros, que en la mejora que la política pública ha generado en el bienestar social.

En primer lugar, la redacción es de baja calidad, y no es homogénea ni en su tono ni en su nivel de concreción. Por un lado, incluye, innecesariamente, definiciones muy poco útiles, además de discutibles, de conceptos como «Dato», «Indicador», «Metodología», «Encargo de evaluación»²⁰. Por otro, al definir en su capítulo III un concepto tan esencial para la Ley como es el de «equipo evaluador», sorprende que se limite a afirmar que: «Las evaluaciones las realizarán equipos externos a los órganos gestores de las políticas evaluadas» (art. 15.2), sin indicar de dónde saldrán tales equipos, qué condiciones deberán cumplir sus miembros, o quien los designará.

La ley se refiere a la evaluación como «mecanismo de lucha contra la inequidad» (Preámbulo IV), pero es evidente que esto debe ser objeto de políticas específicas, no del proceso de evaluación. Asimismo, enfatiza la «participación de los agentes implicados y de los sectores de la ciudadanía afectados. Se promoverá un enfoque inclusivo, que corrija las inequidades sociales» (art. 9.2, 9.3), pero es claro que precisamente los agentes afectados por la política, si bien deben aportar información, no deben participar en la evaluación. La ley hace 9 referencias al plan bienal de evaluación departamental; pero no es lo mismo un plan bienal de evaluación departamental, que sugiere que cada dos años se evaluará el desempeño de un departamento en ciertas áreas por especificar, que un plan departamental de evaluación, que significa un plan de evaluación elaborado por dicho departamento. En todo caso, seguramente haremos lo segundo, aun habiendo aprobado lo primero, sin que ello cause ninguna dificultad.

El concepto de evaluación en la Ley difiere de la descripción que hice al comienzo. En referencia a la evaluación ex-ante se indica que ésta «tendrá por finalidad analizar, con un enfoque transversal y participativo, el valor público de las actuaciones objeto de análisis. A estos efectos son relevantes aspectos como el impacto en la igualdad de género, en materia de reto demográfico, en la transformación digital, en la transición verde, o en las repercusiones en el ámbito de la infancia, la inclusión social o en la racionalidad administrativa, entre otros» (art. 26.3). Las repetidas referencias que se hacen a estos aspectos en el texto de la Ley sugieren que son ellos los que se pretende evaluar²¹, más

²⁰ Algunos aspectos importantes quedan sin especificar, mientras otros son objeto de descripción muy detallada, pero muy poco relevante: «En la selección de herramientas, se dará preferencia a la utilización mixta de técnicas cuantitativas y cualitativas y a la triangulación de datos y resultados.» (art. 20.2). En cuanto a la definición de algunos conceptos, se dice: *Encargo de la evaluación: Es la acción de encomendar la realización de la evaluación de una política pública a un equipo evaluador por parte del órgano responsable de la evaluación de esa política. Dato: Es la representación de una variable cualitativa o cuantitativa, que permite el conocimiento exacto de un hecho o sirve para deducir las consecuencias derivadas del mismo. Indicador: Es un dato o información, individual o agrupada, que permite medir el progreso o efectos de una política pública* (art. 2).

²¹ Por ejemplo: (art. 8.3): *La evaluación tendrá una finalidad y alcance definidos y deberá asentarse en el rigor científico, la objetividad y la solidez metodológica. Con carácter general, atenderá a la perspectiva de género, del equilibrio intergeneracional, del reto demográfico y de la cohesión territorial, de la transfor-*

que el potencial de cumplimiento de los objetivos de una política o norma, o su elección entre otras alternativas. De hecho, esto es lo que vienen transmitiendo las Memorias de Análisis de Impacto Normativo.

La ley también es escasamente ambiciosa acerca del ámbito de la evaluación ex-post. Hay repetidas referencias en la ley a las evaluaciones ex-post de revisión del gasto público que realiza la AIREF. Pero es ésta es la única concepción considerada de la evaluación ex-post, muy lejos de proponer que la evaluación de impacto se aplique con generalidad, en ámbitos como la educación, la sanidad, y el empleo, especialmente en las políticas más ambiciosas por la relevancia de su objetivo social o su presupuesto.

La ley de Institucionalización de la Evaluación deja, por tanto, una excesiva ambigüedad respecto a quién evaluará, así como a los tipos de evaluación que se llevarán a cabo, y, además, excluye explícitamente de su ámbito de aplicación a cualquier actividad de evaluación que ya sea objeto de regulación específica en la normativa existente (art. 3.3): en particular, las que realiza la AGE o las MAIN, en cuya evaluación colaborará la Agencia Estatal de Evaluación de Políticas Públicas. Tampoco se hace referencia a las metodologías estándar de evaluación de políticas públicas, sobre las que existe una amplia literatura. Esta falta de concreción es muy relevante porque la evaluación aparece mencionada en las distintas leyes que se han aprobado posteriormente, pero nuevamente de modo equívoco, sin especificar quién evaluará, cómo lo hará, y cómo se utilizarán sus resultados²².

Una última dificultad surge porque la ley establece que la evaluación se diseñe una vez implementada la política pública objeto de evaluación, sin haber dispuesto a lo largo de su aplicación la recogida de la información estadística precisa, lo cual limita en gran medida las posibilidades de llevar a cabo una evaluación rigurosa. También el presupuesto de evaluación se fija posteriormente a la implementación de la política.

mación digital, de la sostenibilidad social y medioambiental. También: (art. 13.2): *La Agencia Estatal de Evaluación de Políticas Públicas elaborará indicadores generales de evaluación ex ante de la medición efectiva de los resultados esperados en la igualdad de género, en la infancia, en el equilibrio intergeneracional, en el reto demográfico, en la transformación digital, empleo, justicia social y sostenibilidad medioambiental.*

²² Por ejemplo, en la Ley de Empleo, Ley 3/2023, de 28 de febrero: «El seguimiento y evaluación de los resultados de la gestión de las políticas activas de empleo se realizará en el seno de la Conferencia Sectorial de Empleo y Asuntos Laborales y del Consejo General del Sistema Nacional de Empleo, recogién-dose en un Informe Conjunto sobre el empleo que será la base para facilitar el intercambio de buenas prácticas y para definir los instrumentos de planificación y coordinación de la política de empleo a corto y largo plazo.» (art. 11.2). También: «La política de empleo será objeto de una evaluación continuada, planificada, dirigida en el seno del Sistema Nacional de Empleo y realizada con carácter externo e independiente, desarrollándose de manera adicional al conjunto de evaluaciones internas que se realicen dentro del Sistema Nacional de Empleo. Para asegurar su utilidad, la evaluación de la política de empleo perseguirá los objetivos finales que con carácter general contemple la evaluación de políticas públicas en el Estado.» (art. 66.1). Por el contrario, en la Ley de Vivienda no aparece el término «evaluación».

7. VALORACIÓN DE LA LEY

La Ley de Evaluación responde a compromisos adquiridos en los componentes 11 y 29 del Plan de Recuperación. Muchas de las actuaciones de dicho plan se incluyeron a requerimiento de la Comisión Europea; entre ellas, se pedía a España un impulso a la evaluación de políticas públicas. Curiosamente, sin embargo, el propio Plan de Recuperación no incluye ningún sistema de evaluación del impacto de sus múltiples actuaciones. Tan solo manifiesta una clara preocupación por supervisar el rigor en la ejecución contable-financiera que asigna al Tribunal de Cuentas y la Intervención General del Estado, reduciendo así, de nuevo, la evaluación económica al control contable.

La Ley de Evaluación se tramitó por procedimiento de urgencia (BOCG-C, 2022), aduciendo la cercanía de la fecha de liberación de un nuevo tramo de los fondos Next Generation para España; al igual que está sucediendo con muchas de las leyes que se están aprobando desde que se constituyeron dichos fondos europeos²³. Al acortar los tiempos de consulta pública, el trámite de urgencia dificulta la recepción de sugerencias y propuestas. Quizá por esta premura, la Memoria de Impacto Normativo de la Ley de Evaluación no contiene ninguna información sobre un tema tan relevante como las experiencias de otros países en términos de sus estructuras y procesos de evaluación.

Tampoco hay referencias, probablemente porque no se produjeron, acerca de las consultas establecidas con los muchos expertos que, tanto en el mundo académico, como en diversas instituciones de investigación, trabajan en los temas relacionados con la evaluación de políticas públicas²⁴. Esto, a pesar de que en los dos últimos años, tres de las mejores publicaciones académicas españolas en Economía: Papeles de Economía Española (Funcas), Información Comercial Española (Ministerio de Industria, Comercio y Turismo), y Economía, han dedicado números monográficos a la evaluación de políticas públicas, con cuyo material podría haberse diseñado una Ley bastante más acertada. Llama también la atención que, como consta en su Memoria de Impacto Normativo, ministerios responsables de áreas en las que la evaluación de políticas es de la mayor importancia, como son Educación, Industria, Trabajo, Sanidad, no presentasen observaciones previas a la reunión de la Comisión Delegada para AA. EE.

²³ Entre ellas, la Ley de Vivienda, la Ley de Empleo (Ley 3/2023), la Ley de Función Pública, la Ley de Garantía Integral de Libertad Sexual (Ley Orgánica 10/2022), Ley de Cooperación para el Desarrollo Sostenible y la Solidaridad Global (Ley 1/2023), Ley de Protección de los Derechos y Bienestar de los Animales (Ley 7/2023), Ley del Sistema Universitario (Ley Orgánica 2/2023).

²⁴ Lamentablemente, está siendo una tradición que las MAIN no cumplan algunos de estos requisitos. Tampoco es sencillo encontrarlas, pues no existe un depósito común para ellas, y no queda sino utilizar un buscador con el nombre de la ley especificando que se desea acceder a su MAIN.

Lamentablemente, la Ley de Institucionalización de la Evaluación de Políticas Públicas no genera mucho optimismo respecto a avances significativos en tan importante esfera de la gestión pública. Se ha optado por recuperar la antigua AEVAL sin abordar los tres problemas que causaron su fracaso: una reducida dotación presupuestaria y de personal, dificultad para acceder a personal con los conocimientos técnicos necesarios y la ausencia de incentivos para someter las iniciativas políticas a un análisis riguroso por parte de evaluadores que estarán excesivamente cercanos a los organismos evaluados.

Se crea con la ley un conjunto de estructuras administrativas, coordinadas desde la Secretaría de Estado de Función Pública, sin justificar su utilidad. Como viene sucediendo desde hace algún tiempo, da la impresión de que existe una mayor preocupación por aprobar normas y crear nuevas instituciones que por dotarlas de un contenido riguroso y un sistema de seguimiento del cumplimiento de sus objetivos²⁵. Conviene de todos modos observar que, literalmente, la ley únicamente «se autoriza la creación de la Agencia Estatal de Evaluación de Políticas Públicas» (art. 29), por lo que también existe la posibilidad de que no llegue a crearse dicha agencia y la ley quede sin aplicación.

A falta de alguna norma posterior que desarrolle y matice los aspectos que he descrito, el texto aprobado no permite confiar que la nueva generación de informes de evaluación vaya a ser de mejor calidad que los ya existentes, que tienen escaso valor añadido.

8. EXPERIENCIA INTERNACIONAL EN MATERIA DE EVALUACIÓN DE POLÍTICAS PÚBLICAS

En algunos países europeos, Alemania, Francia, Holanda, y especialmente el Reino Unido, así como en Canadá y EEUU, la evaluación de políticas públicas se ha ido desarrollando desde los años 60 y 70 del pasado siglo y, en ellos, actualmente sería inaceptable no evaluar las principales actuaciones públicas.

En OCDE (2020) se examina las prácticas de evaluación de políticas públicas en 42 países, mostrando que, al igual que en España, en muchos países la evaluación era más un compromiso pendiente que una realidad. Los países de la OCDE incluidos en este estudio suelen tener una base legal para orientar la política de evaluación, siendo menos frecuente en el resto. Sus experiencias son relativamente heterogéneas: la responsabilidad de evaluación

²⁵ Así lo sugiere la avalancha de Planes, Estrategias y leyes aprobadas recientemente, muchas de las cuales se solapan si que se haya establecido una coordinación ni entre sus objetivos ni en su aplicación.

dentro del gobierno se centra en el Ministerio de Finanzas/Hacienda o en algún centro dependiente de Presidencia. Fuera del ejecutivo, suelen existir Supreme Audit Institutions (SAI), generalmente de carácter independiente y profesional, que son entidades públicas cuya función de auditoría externa está establecida por la constitución o el órgano legislativo supremo. En estos casos, está bien diferenciada la evaluación de impacto de la supervisión financiero-contable²⁶.

Estos organismos suelen estar bien coordinados con el resto de las agencias que llevan a cabo evaluaciones dentro de la administración, dentro y fuera de los ministerios. En general, existe una preocupación explícita acerca de que las distintas agencias que realizan evaluación tengan independencia del ejecutivo. Es habitual contar con centros de investigación y universidades para colaborar en las tareas de evaluación, aunque también es frecuente el recurso a consultoras privadas. Un breve resumen de sistemas internacionales de evaluación se recoge en el Apéndice 2.

Por otra parte, los análisis comparados realizados por la UE entre sus miembros muestran que algunas de las limitaciones de nuestra tradición en evaluación de políticas públicas son compartidas con otros países: a) inadecuados presupuestos de evaluación, b) desinterés por la generación de la información precisa para evaluar, c) ausencia de un marco legal adecuado, d) la falta de calidad de la evaluación, e) la falta de interés político por la evaluación, f) el escaso uso de la evaluación para la mejora del diseño de futuras actuaciones y regulaciones. Pero imagino que no estaremos dispuestos a admitir una incapacidad relativamente extendida como excusa para limitar nuestro progreso institucional.

Entonces,

9. ¿QUÉ DEBERÍA CAMBIAR EN EL MARCO REGULATORIO PARA ESPAÑA DE LA EVALUACIÓN DE POLÍTICAS PÚBLICAS?

Los análisis comparados realizados por reconocidos investigadores internacionales [Blume y Voight (2011) y (2013), Haiarinejad *et al.* (2012)] sugieren cinco factores que condicionan el éxito de un sistema de evaluación de políticas públicas: la independencia de la entidad evaluadora respecto del po-

²⁶ Las Supreme Audit Institutions (SAI) realizan generalmente una supervisión del gasto público, que sigue siendo una parte central de la cartera de auditoría. Las SAI realizan: auditorías financieras para evaluar la fiabilidad y precisión de los informes financieros de las entidades públicas, y auditorías de cumplimiento, para evaluar el cumplimiento de las entidades públicas de los requisitos legislativos normativos que las definen. Recientemente se observa que el papel de las SAI está evolucionando, ya que adoptan una visión cada vez más amplia e integral de la fiabilidad, eficacia y eficiencia de las políticas y los programas. Sin duda, las SAI tienen un potencial sin explotar para ir más allá de su rol tradicional de supervisión y aportar evidencia para una formulación de políticas más informada.

der ejecutivo; su adscripción al Parlamento; el rigor en la formación de su personal y la independencia en la selección y nombramiento de sus responsables; la transparencia de los resultados de las evaluaciones; y su relevancia en la rendición de cuentas de los gestores públicos, dada la importancia que ésta tiene en una democracia sana.

Avanzar en la línea de los cinco factores de éxito que acabo de mencionar, y reducir o eliminar sus actuales limitaciones, lo que no se consigue con la reciente Ley de Evaluación, son los dos requisitos para mejorar nuestro sistema de evaluación de políticas públicas. Para ello, sugiero cinco propuestas:

1. La Agencia Nacional de Evaluación de Políticas Públicas debe tener independencia efectiva respecto de la administración, con autonomía para establecer su programa de trabajo y para elaborar su presupuesto, así como mecanismos flexibles de contratación laboral fuera de la administración.
2. Es importante aprovechar la infraestructura existente en universidades y organismos públicos de investigación que cuentan con buenos expertos en métodos y prácticas de evaluación. La colaboración con Universidades podría extenderse a la identificación de las competencias necesarias para el desempeño de la función evaluadora y el diseño de planes específicos de formación en materia de evaluación para empleados públicos²⁷, tareas que la Ley considera únicamente al Instituto Nacional de Administración Pública y al Instituto de Estudios Fiscales.
3. La posible dependencia respecto del Parlamento de la Agencia de Evaluación, quizá a través de la Oficina Presupuestaria de las Cortes Generales, garantizaría la mayor transparencia en los resultados de sus análisis de evaluación y dificultaría la posible captura del evaluador por parte de las agencias evaluadas. Si realmente existe una decidida vocación de evaluación por parte de los distintos gobiernos, debería manifestarse en una interacción fluida entre la Agencia de Evaluación y una unidad, posiblemente dependiente de Presidencia de Gobierno o del Ministerio de Hacienda, con la cual hacer un seguimiento de las evaluaciones y traducir sus resultados en reformas.
4. Reformar cinco aspectos del funcionamiento de la administración: Homogeneizar los procedimientos de recogida de información de las AAPP. Mejorar la colaboración entre las Administraciones que

²⁷ Los conceptos y metodologías de evaluación apenas están presentes en las titulaciones actuales, pero esto puede resolverse mediante la creación de nuevas titulaciones o programas de especialización.

son propietarias de la información y las entidades que podrían llevar a cabo evaluaciones de impacto. Eliminar las dificultades de coordinación entre agencias autonómicas (por ejemplo, dentro del Sistema Nacional de empleo, como apuntó la AiReF en su evaluación de 2019). Establecer enlaces entre las bases de datos de las distintas agencias de la administración pública (Seguridad Social y Servicio Público de Empleo, pero también, Agencia Tributaria, INE). Evitar una aplicación excesivamente estricta, a veces muy restrictiva, de la Ley de Protección de Datos.

5. Como parte de la evaluación ex-ante, las Memorias de impacto normativo de las nuevas leyes deberían someterse a revisión de la Agencia de Evaluación, con la capacidad de devolverlas, en caso de deficiencias, a los departamentos promotores.
6. Toda política pública debe especificar, como parte de su diseño, el procedimiento de evaluación que se utilizará, su dotación presupuestaria y la recogida de la información precisa. El gobierno debe responder de los resultados finales, en relación con las expectativas iniciales y el presupuesto ejecutado, y los resultados de las evaluaciones deben hacerse públicos y ser fácilmente accesibles a todos los ciudadanos, a través de un repositorio de evaluaciones.

10. ¿HACIA DÓNDE VA LA EVALUACIÓN DE POLÍTICAS PÚBLICAS?

En las últimas décadas, la evaluación de políticas públicas ha pasado a ocupar un papel principal en la regulación de organismos internacionales, como la OCDE y la UE, a la vez que ha captado el interés de amplios grupos de académicos e investigadores, además del reconocimiento internacional, culminado en la concesión de los premios Nobel en Economía de 2019 y 2021 a que antes hice referencia. La UE ha avanzado considerablemente en el desarrollo de sus procesos de evaluación, que exige crecientemente a los receptores de sus programas de financiación. Insiste en utilizar evaluaciones basadas en evidencia, y enlazar la evaluación ex-ante y ex-post de sus programas, porque la política pública debe verse como un proceso continuo.

Se ha producido un cambio de paradigma hacia lo que denomina la «nueva gobernanza de lo público», que sitúa a la sociedad civil en el centro de la gestión pública. Los ciudadanos no son simples votantes y consumidores, sino que participan en la resolución de problemas y en la creación de bienes públicos. El papel del gobierno consiste entonces en «ordenar y facilitar los procesos de interacción de modo que se facilite la participación ciudadana y sus intereses se articulen y traten de manera abierta, transparente y equilibrada»

[Kubera (2017)]. Desde este punto de vista de la gestión pública, el diálogo y la interacción entre las partes interesadas en el curso del proceso de evaluación pueden mejorar el capital relacional, que es crucial en el nuevo paradigma de gobernanza pública.

A su vez, la UE ha ido extendiendo sus actividades de evaluación desde los programas de gasto hasta las actuaciones regulatorias, y su agenda por una Mejor Regulación (*EU's Better Regulation agenda*) trata de diseñar y evaluar las políticas y la legislación comunitaria de modo abierto y transparente, de modo que los ciudadanos y las partes interesadas tengan la oportunidad de contribuir a lo largo de todo el ciclo político y del proceso de elaboración de leyes²⁸. En este enfoque encajan perfectamente los procesos de consulta pública y de información pública que deben acompañar a todo anteproyecto de ley [de acuerdo con la Ley del Gobierno (Ley 50/1997)] y que deben ser vistos como parte de la evaluación ex-ante de una determinada regulación o actuación pública.

Esta concepción participativa de la elaboración normativa motiva también la Alianza para un Gobierno Abierto, en la que participa España desde su creación en 2011, y dentro de la cual hemos suscrito nuestro IV Plan de Gobierno Abierto para el periodo 2020-2024. El concepto de «gobierno abierto» incorpora tres elementos fundamentales: transparencia, accesibilidad para cualquier persona y en cualquier lugar, y capacidad de respuesta a las nuevas ideas y demandas. Incorpora de manera clara la idea de dialogar e interactuar con el ciudadano, haciéndole partícipe activo de la formulación de las políticas públicas y decisiones que impactan sus vidas, no un mero receptor pasivo de políticas y normativa²⁹.

11. REFLEXIONES FINALES

La concepción participativa de la gobernanza abre un escenario prometedor en refuerzo de la calidad democrática, en el que la evaluación de la puesta en práctica y el impacto de las políticas públicas juegan un papel central. Los responsables públicos deben dejar de percibir la evaluación externa como un mero enjuiciamiento de su gestión, y reconocer la necesidad de rendir

²⁸ La *Better Regulation Agenda* de la Comisión Europea fomenta la aplicación de principios y recomendaciones en materia de: consultas públicas, evidencias empíricas, evaluaciones de impacto inicial, simplificación de normas («one in, one out»), reducción de burocracia, incorporación de los Objetivos de Desarrollo Sostenible, y previsión estratégica en medioambiente, digitalización, y contexto socioeconómico y geopolítico. Las actuaciones en materia de evaluación de la UE pueden consultarse en *EU evaluation reports*

²⁹ Todo ello, sobre la base de los principios de buen gobierno, tales como la rendición de cuentas, justicia, equidad, eficiencia, respeto del imperio de la ley y altos estándares de comportamiento ético (*Open Government Partnership*, <https://www.opengovpartnership.org/estrategia-de-ogp-2023-2028/>).

cuentas a la sociedad, basada en evidencia observable. La evaluación debe verse como un aprendizaje que permite evitar ineficiencias futuras y mejorar la eficacia de la acción de gobierno. Debe realizarse desde las agencias públicas constituidas a tal efecto, pero también desde el mundo académico y otras instancias de la sociedad civil. Para ello, los investigadores deben poder acceder, en formato fácilmente utilizable, a los datos necesarios para llevar a cabo su análisis de evaluación³⁰. El debate que pueda suscitarse sobre los métodos y los resultados de la evaluación garantizará la transparencia en el acceso a la información, y redundará en un mejor diseño y puesta en práctica de las políticas públicas, siendo beneficioso para la sociedad.

Pero una nueva normativa, incluso si está correctamente diseñada, no garantiza por sí sola que se instaure una estructura rigurosa de evaluación de políticas públicas en beneficio de todos los ciudadanos. Cuatro condiciones parecen esenciales para el éxito de dicho sistema: 1) la existencia de una cultura social de la evaluación, de acuerdo con la cual la sociedad demanda que se evalúen las actuaciones públicas, y concede suficiente importancia a la formación de especialistas; 2) un hábito de interacción, en cada ámbito de aplicación de una política o una nueva normativa, con los expertos, la sociedad civil, las universidades y centros de investigación; 3) un interés por la comunicación con los ciudadanos, como sujetos activos en la formulación de las políticas públicas; y, muy especialmente, 4) una firme voluntad política a favor de la evaluación, facilitando su realización, y utilizando sus resultados para mejorar el diseño de políticas y para una rigurosa rendición de cuentas. Estas son las líneas en las que hemos de avanzar.

BIBLIOGRAFÍA

- AIREF (2020): «Opinión para una estrategia de acceso a datos administrativos», *Opinión* 1/20.
- (2022): «1.ª Opinión. Ingreso Mínimo Vital», *Opinión* 1/22.
- ALBI, E., y ONRUBIA, J. (2016): «Institucionalizar la evaluación de políticas públicas: eficiencia y rentabilidad social», *Papeles de Economía Española*, núm. 147, pp. 50-68.
- (2018): *Supervisión, control y evaluación para una gestión pública transparente*, Instituto de Estudios Fiscales, Madrid.
- ALMUNIA, M., y REY BIEL, P. (2021): «La gestión de los datos administrativos en España: diagnóstico y retos de futuro», *Cuadernos Económicos de ICE*, 102, pp. 103-133.

³⁰ «La AIREF (2020) propuso el diseño de una estrategia nacional (30) que fue compartida también por la comunidad científica e investigadora (Almunia y Rey Biel, 2021) y que debería requerir avances legales para que los datos se produzcan, además de con fines estadísticos, con fines de investigación y evaluación de políticas públicas a la vez que se crean uno o varios repositorios de la información que sean accesibles por las instituciones y la academia.» (Casado y Del Pino, 2022, p. 11).

- ANGRIST, J., y PISCHKE, J. S. (2009): *Mostly Harmless Econometrics: An Empiricist's Companion*, Princeton University Press.
- BLUME, L., y VOIGHT, S. (2011): «Does organizational design of supreme audit institutions matter? A cross-country assessment», *European Journal of Political Economy*, 27, pp. 215-229.
- (2013): «The economic effects of constitutional budget institutions», *European Journal of Political Economy*, 29, pp. 236-251.
- CASADO, J. M., y DEL PINO, E. (2022): «Similitudes y diferencias en la evaluación de políticas públicas en ocho países: España en perspectiva comparada», *Papeles de Economía Española*, núm. 172, pp. 2-18.
- CEBRÍAN, I.; CUETO, B., y MATO, F. J. (2022): «La evaluación de impacto de las políticas activas de empleo en España: el deseo y la realidad», VV.AA. (2021), *Información Comercial Española*, pp. 133-157.
- CONGRESO DE LOS DIPUTADOS: Boletín Oficial de las Cortes Generales (BOCG-C, 2022) *Proyecto de Ley de institucionalización de la evaluación de políticas públicas en la Administración General del Estado*.
https://www.congreso.es/public_oficiales/L14/CONG/BOCG/A/BOCG-14-A-103-1.PDF.
- FUENTE, A. DE LA (COORDINADOR); RUS, G. DE; FERNÁNDEZ, M.; GARCÍA, M. A.; JANSEN, M.; JIMÉNEZ, S.; NOVALES, A.; ONRUBIA, J.; PÉREZ RENOVALES, J.; SASTRE, E., y SICILIA, J. (2021): «La evaluación de políticas públicas en España: antecedentes, situación actual y propuestas para una reforma», *Informes y papeles del Grupo de Trabajo Mixto Covid-19. FEDEA Policy Paper*, 2021-09, Madrid.
<https://bit.ly/3mnb4Y5>.
- FUENTE, A. DE LA, y NOVALES, A. (2023): «El marco institucional para la evaluación de políticas públicas en España», *Ekonomiaz*.
- FUENTE, A. DE LA (2023): «Sobre las memorias de impacto de las leyes: un mal ejemplo y una propuesta», blog FEDEA.
<https://policy.fedea.net/sobre-las-memorias-de-impacto-de-las-leyes-un-mal-ejemplo-y-una-propuesta/>.
- GOBIERNO DE ESPAÑA (2021): *Plan de Recuperación, Transformación y Resiliencia, España Puede*, 16 de junio de 2021.
- HAIARINEJAD, G.; SHEKARBEGI, S.; KAZEMI, A. A., y JAMELI, S. (2012): «Supreme audit court of auditors' insights on operational audit challenges», *Management Science Letters* 2 (2012), pp. 757-762.
- INSTITUTO PARA LA EVALUACIÓN DE POLÍTICAS PÚBLICAS (2021): *Estudio de Diagnóstico de la Evaluación de Políticas Públicas en la Administración General del Estado*.
- KUBERA, P. (2017): «A new approach to policy evaluation in the European Union», *Przegląd Polityczny*, 4, pp. 21-30.
- MATHOT, A., y GIANNINI, F. (2022): «Evaluation Framework and Practices: A comparative analysis of five OECD countries», *Journal on Budgeting*, volume 2022 issue 2, OECD.
- OECD (2020): *Improving Governance with Policy Evaluation: Lessons From Country Experiences*, OECD Public Governance Reviews, OECD Publishing, Paris.
<https://doi.org/10.1787/89b1577d-en>.

- ROBBINS, L. (1932): *An Essay on the Nature and Significance of Economic Science*, Mac Millan, London.
- RUS, G. (2023): «El valor social de la evaluación económica de las políticas públicas», *Ekonomiaz*, núm. 103.
- SEBASTIÁN, C. (2019): *Para que España avance*, Galaxia Gutenberg.
- STOCKMANN, R.; MEYER, W., y TAUBE, L. (eds.) (2020): *The Institutionalization of Evaluation in Europe*, Cham: Palgrave Macmillan.
- VV.AA. (2004): *Encyclopedia of Evaluation*, SAGE Publications.
- VV.AA. (2021): «La evaluación de las políticas públicas en España: recursos, metodologías y algunos ejemplos», *Cuadernos Económicos del ICE*, núm. 102, coordinado por Juan Francisco Jimeno.
- VV.AA. (2022): «Evaluación de políticas públicas», *Papeles de Economía Española*, núm. 172, coordinado por Francisco Pedraja Chaparro.
- VV.AA. (2023): «Calidad de las finanzas públicas: el papel de la evaluación», *Ekonomiaz*, núm. 103.

APÉNDICE 1: METODOLOGÍAS DE EVALUACIÓN

El ideal de evaluar una política consistiría en comparar el nivel de bienestar social con y sin la actuación pública. Lamentablemente, este es un deseo irrealizable, pues el bienestar social no es observable y, por el Teorema de Imposibilidad de K. Arrow, tampoco podemos agregar preferencias individuales, supuesto que pudiésemos averiguarlas, en un indicador social.

Durante décadas, la metodología más extendida, especialmente en la evaluación de infraestructuras, ha sido la que se conoce como análisis Coste-Beneficio, que ha tratado de relacionar el beneficio social generado por la actuación pública, con su coste. Esta metodología se enfrenta a tres dificultades esenciales: una de ellas es que, si bien es sencillo medir costes, es mucho más difícil beneficios, puesto que muchos son intangibles, no cuantificables en términos económicos. En ocasiones, se resuelve esta dificultad preguntando a los potenciales afectados por la intervención pública cuánto estarían dispuestos a pagar por disfrutar de los beneficios esperados de la misma. Este enfoque puede ser aplicable en políticas que tratan de implantar o mejorar un determinado servicio de transporte o comunicaciones, o en determinadas políticas de educación o sanidad, pero puede ser menos apropiada en otros contextos. Otra dificultad surge porque, generalmente, los efectos de una política se dejan notar durante un periodo de tiempo que puede ser dilatado, por lo que es preciso aplicar una tasa de descuento a los efectos futuros, para hacerlos comparables con los efectos más inmediatos. Pero no conocemos cuál es la tasa de descuento apropiada y, además, variaciones pequeñas en ella pueden implicar grandes cambios en la contabilidad de costes y beneficios. Una tercera dificultad proviene de la incertidumbre asociada a los efectos de una política; su tratamiento debería hacernos considerar varios escenarios posibles, y asignar a cada uno de ellos una determinada probabilidad de ocurrencia, para llevar a cabo la comparación de costes y beneficios. Estas dificultades, que son en buena medida inherentes al análisis económico, requieren un cuidadoso tratamiento.

La no observabilidad de la utilidad y el bienestar hace que frecuentemente, la evaluación se simplifique, sustituyendo los beneficios sociales de un proyecto por medidas de la efectividad del gasto. La evaluación se reduce entonces a analizar el coste que ha representado conseguir una unidad del objetivo, ya sea un km de vía férrea o autovía, la graduación de un estudiante adicional, la reducción de un día en el tiempo de espera para determinada intervención quirúrgica, o a consultas en la sanidad pública, etc.. En ocasiones se aplican criterios multiobjetivo, porque la política se introdujo con varios objetivos potenciales. Pero es delicado tratar de consolidar criterios que pueden ser muy heterogéneos en un único indicador, que pudiera permitir decidir en positivo o negativo si implementar la política. En todo caso, este sería un análisis

sis de efectividad de costes, pero no una evaluación propiamente dicha, pues no pretende medir las ganancias de bienestar³¹.

Otro grupo de metodologías trata de medir la eficiencia productiva, que consiste en evaluar si, con las cantidades de inputs utilizadas en un proceso productivo podrían obtenerse más unidades del output o si para obtener las unidades de output observadas, podrían utilizarse menos unidades de los inputs utilizados en el proceso. Este enfoque es útil para comparar la eficiencia productiva entre unidades homogéneas productoras de un mismo servicio público, permitiendo establecer rankings entre unidades, e identificando las causas de su posible ineficiencia. Entre los métodos no paramétricos de este tipo se encuentra el Análisis Envolvente de Datos. Alternativamente, se utilizan los métodos de estimación econométrica de fronteras eficientes, pudiendo valorar la distancia de cada unidad productiva a la frontera eficiente. En este enfoque, el Índice de Malquist permite medir la evolución temporal de la eficiencia de una determinada unidad productiva, descomponiendo los cambios observados en productividad según obedezcan a cambios en el progreso técnico o a mejoras de eficiencia.

Para la evaluación de impacto, que tiene carácter ex-post, ha cobrado gran relevancia en los últimos años la utilización de experimentos, que permiten estimar el impacto causal mediante comparación de muestras de control y de tratamiento. Por comparación con las ciencias experimentales, se trata de comparar los resultados obtenidos en unidades en las que se ha aplicado una determinada política, que forman el grupo de tratamiento, con otras en no sometidas a la misma, que forman el grupo de control. Afortunadamente, además, este es un área donde se han producido recientemente considerables avances que han motivado, en parte, la concesión del Premios Nobel en Economía a A. Banerjee y E. Duflo en 2019 y a D. Card, J. Angrist, y G. Imbens en 2021. Del enfoque experimental en Economía nos habló nuestra compañera Mari Paz Espinosa en su discurso de ingreso en esta Real Academia.

La experimentación como enfoque de evaluación de políticas está sujeto a validación interna y externa. La *validación interna* requiere que no exista *sesgo de selección*, es decir, que no se ha seguido ningún criterio concreto para asignar cada observación muestral a un grupo o al otro; esencialmente, significa que los grupos de control y tratamiento son comparables. La comparabilidad puede presentarse de dos formas: en algunas ocasiones es posible utilizar unidades de observación que sean totalmente homogéneas en aquellos aspectos que pueden condicionar los resultados de la política; en otras, las uni-

³¹ La autovía podría conducir a una zona muy poca poblada, con un grado de utilización muy reducido. De poco importaría entonces que la construcción haya sido muy eficiente en costes, si pocas personas se benefician de la misma; la evaluación de la obra debería ser negativa, no positiva.

dades de observación son muy diversas, pero se decide donde aplicar la política de modo puramente *aleatorio*, garantizando que existe una diversidad comparable en el grupo de control y el grupo de tratamiento. Por ejemplo, para evaluar el efecto que sobre los resultados escolares tiene una política que impone un tamaño máximo de estudiantes por grupo, compararíamos los resultados escolares obtenidos en centros en los que se ha impuesto un tamaño máximo de clase, con otros en los que no se ha impuesto. Es esencial que unos y otros centros no difieran sistemáticamente en aquellas características que pueden condicionar los resultados escolares, además del tamaño de la clase, como pudiera ser la renta familiar, u otros. Es decir, si se decide dónde aplicar la política de modo puramente aleatorio, debe existir una diversidad comparable en ambos grupos de centros, el grupo de control y el grupo de tratamiento. La *validación externa* requiere que los resultados de la experimentación realizada sean aplicables a otros contextos, si se quiere presentar con carácter general dichos resultados. Cuando esto no es posible porque existen diferencias apreciables en las características que condicionan los efectos de la política, hay que limitar el resultado de la evaluación al contexto en que se llevó a cabo.

Endogeneidad y omisión de efectos causales

Una cuestión central en el análisis de efectos causales se refiere a la posible omisión de otras causas en la explicación de un determinado fenómeno, como los resultados escolares de un grupo de estudiantes, o el salario de un conjunto de trabajadores; en tal caso, podemos asociar, equivocadamente, a la causa que consideramos el agregado de su efecto directo, y el de las causas adicionales omitidas del modelo. Este sería el caso si estimamos el efecto que sobre el salario de un trabajador tiene el número de años de escolarización, sin tener en cuenta que sus capacidades naturales (medidas por algún indicador de inteligencia, habilidades manuales, etc.) también cabe esperar que influya sobre su salario. Establecer una relación causal entre la asistencia a la educación preescolar y los resultados educativos, intentando corregir posibles sesgos provocados por la endogeneidad implícita de esta decisión (los años de asistencia a educación infantil están correlacionados con variables no observables como sucede con la motivación de los padres). (Santín y Sicilia, 2015).

Otro tipo de situaciones problemáticas es el conocido como causalidad inversa; por ejemplo, si analizamos el efecto de los ingresos como determinantes del grado de felicidad de un trabajador, debemos considerar que un trabajador que se siente más feliz, puede trabajar mejor y tener un salario más elevado. Si analizamos el efecto de la repetición de curso sobre los resultados escolares, no podemos ignorar que son precisamente los malos resultados escolares los que conducen a la repetición de curso. Ambas situaciones, la omisión de factores relevantes, y la causalidad inversa aparecen en muchos de los

análisis empíricos en ciencias sociales, y hay que tratarlos adecuadamente, pues ambos generan sesgos en la estimación de la magnitud de efectos causales.

Para poder evaluar correctamente los efectos de las intervenciones públicas es necesario corregir los posibles sesgos que provoca la existencia de endogeneidad en los datos, derivada de la omisión o imposibilidad de observación de factores relevantes, la existencia de causalidad inversa, el sesgo de autoselección o los errores de medición en las variables de interés observadas. La búsqueda de soluciones para evitar los problemas vinculados con la endogeneidad se ha convertido en una de las principales preocupaciones de los investigadores que trabajan con experimentación en economía.

La mejor manera de corregir los posibles sesgos de selección e identificar los efectos causales es mediante la realización de un experimento aleatorio, es decir, seleccionando aleatoriamente una muestra de la población y, dentro de ella, asignando al azar a los individuos a los grupos de tratamiento y control. Alternativamente, pueden utilizarse aproximaciones cuasiexperimentales, como el método de diferencias en diferencias, la regresión en discontinuidad y el enfoque de variables instrumentales [ver Angrist y Pischke (2009)].

La utilización del enfoque de *diferencias en diferencias* requiere disponer de información sobre los individuos que componen los grupos de tratamiento y control antes y después de la implementación de una determinada intervención, lo que requiere disponer de observaciones de dichos individuos a lo largo de un período de tiempo. Esa metodología resulta especialmente adecuada cuando el tratamiento se aplica a un grupo de individuos, pero no a otros a partir de un determinado momento del tiempo. Normalmente, estas políticas se aplican al grupo de tratamiento a partir de un momento del tiempo, por lo que se pueden observar las diferencias en la variable objetivo entre el grupo tratado y el grupo de no tratados antes y después de la implementación de la política pública y comparar ambas diferencias. Esta doble diferencia permite eliminar el efecto de otros factores, si suponemos que dichos efectos son constantes a lo largo del tiempo.

La aplicación de la *regresión en discontinuidad* se aplica cuando la selección del grupo de tratamiento se basa en un umbral de un determinado indicador, y utiliza únicamente las observaciones muestrales para las que dicho indicador toma valores próximos a dicho umbral, por encima o por debajo del mismo. Por tanto, está condicionado por la disponibilidad de un volumen suficiente de datos alrededor del punto de corte en la variable que determina quiénes podrán beneficiarse del programa y quiénes formarán parte del grupo de control, como suele ser el caso de los datos administrativos. Un ejemplo es el de un programa de becas al que acceden aquellos estudiantes que obtuvieran una nota superior a un valor obtenido mediante un test específico diseñado para el programa. Su validez se basa en que las observaciones por encima y por debajo

del umbral constituyan dos submuestras comparables, de modo que podamos suponer que su asignación al grupo de tratamiento o al grupo de control es prácticamente aleatoria; debemos excluir, por tanto, la existencia de diferencias entre ambos grupos en otras variables distintas a la variable objetivo.

El uso del método de *variables instrumentales* tiene el propósito de eliminar, o al menos mitigar, los problemas de sesgo relacionados con la omisión de variables relevantes o la posible causalidad inversa. Esta alternativa no es tan exigente con los datos necesarios para su aplicación, pero es crucial identificar el instrumento apropiado. Este debe guardar relación con el factor para el cual actúa como instrumento (aquél cuyo efecto se estimaría con sesgo, de no utilizar un procedimiento adecuado). Además, debe relacionarse con el fenómeno que se pretende explicar únicamente a través de su relación con dicho factor; es decir, debe estar incorrelacionado con cualquier otro determinante del fenómeno que se estudia. Por ejemplo, Artés (2014) analiza a qué partido político favorece la abstención, estimando un modelo que explica el porcentaje de voto a distintos partidos en cada municipio en distintas elecciones generales, en función del porcentaje de abstención. El instrumento utilizado es el porcentaje de lluvia en cada municipio en el día de las elecciones; este instrumento afecta a la abstención, pero no al sentido del voto, puesto que las personas que acuden a votar no deciden a quién votar en función de si llueve o no ese día.

Este breve Apéndice no es sino una simple aproximación a los métodos estadísticos y econométricos para una evaluación de impacto. El lector interesado en ampliar el tema puede consultar VV.AA. (2021), VV.AA. (2022), Angrist, J.; Pischke, J. S. (2009).

APÉNDICE 2: LA EXPERIENCIA INTERNACIONAL EN LA EVALUACIÓN DE POLÍTICAS PÚBLICAS³²

Aunque algunos países comenzaron a evaluar sus políticas públicas con anterioridad, la evaluación cobró un especial vigor en la década de los 80, en buena parte por el esfuerzo de organismos internacionales (UE, OCDE, Banco Mundial) por promover estándares homogéneos de evaluación. A partir de 1988, la Comisión Europea comenzó a requerir evaluaciones a los estados receptores de fondos estructurales en las áreas de aplicación de dichos fondos y, a comienzos de este siglo, introdujo la evaluación de impacto regulatorio. También el enfoque de evaluación basado en las ciencias del comportamiento

³² Esta sección se basa en CASADO y DEL PINO (2022). Para un análisis más detallado, el lector interesado puede consultar MATHOT y GIANNINI (2022) y OECD (2020).

ha centrado el interés de la UE, habiendo establecido *nudge centers*³³ en 10 países, entre los que no está España. Las experiencias de algunos de los países son relativamente heterogéneas: en algunos casos, la principal agencia de evaluación está adscrita al Parlamento y, en otros, depende de Presidencia del Gobierno. Estos organismos están bien coordinados con el resto de las agencias que llevan a cabo evaluaciones dentro de la administración; en general, existe preocupación explícita acerca de que las distintas agencias que realizan evaluación tengan independencia del ejecutivo. Suelen colaborar en las tareas de evaluación con centros de investigación y universidades, aunque también es frecuente el recurso a consultoras privadas. Crecientemente se están utilizando en la evaluación elementos de ciencias del comportamiento.

En Suiza, el Parlamento federal cuenta desde 1990 con su propia unidad de evaluación interna, denominada Control Parlamentario de la Administración. Su Tribunal de Cuentas también hace análisis coste-beneficio de políticas públicas, aunque la mayor parte de las evaluaciones se contratan externamente.

La Ley de Cuentas del Estado en Holanda exige que todas las propuestas de política pública incluyan claramente sus objetivos y los instrumentos que usarán para conseguirlos y, en políticas de infraestructuras, la evaluación *ex ante* es obligatoria. El Tribunal de Cuentas también analiza la eficacia y la eficiencia de las políticas y el Parlamento puede determinar si las propuestas que llegan del Gobierno están bien fundamentadas y si las evaluaciones son suficientes. Desde 2006, los ministros son responsables de realizar policy reviews en períodos entre cuatro y siete años, y cada ministerio tiene su propio equipo de análisis del comportamiento, pero no hay unidades de evaluación como tales, y la mayoría se contratan con consultoras, con la academia o con institutos de investigación.

Canadá reguló en 2001 la obligatoriedad de la evaluación en los ministerios, con equipos de seguimiento de dichas evaluaciones, subordinada al Consejo del Tesoro, que coordina la política de evaluación. Desde 2015 existe una Unidad de Impacto e Innovación, dependiente del primer ministro, y una Oficina de Presupuestos en el Parlamento, que proporciona a los diputados análisis independientes sobre el presupuesto o sobre otros temas relacionados con la economía nacional. Por último, la Oficina del Controlador General de Canadá supervisa las actividades del Gobierno y evalúa el gasto público. Realiza una importante labor de auditoría de distintos tipos, incluyendo en algunas de ellas evaluaciones del impacto de los programas públicos de diferentes sec-

³³ El concepto del «nudging» (un término acuñado por el premio Nobel Richard Thaler) es la manera de «empujar» a la población a tomar decisiones que las benefician a largo plazo, algo que tiene influencia tanto en la economía como en la política.

tores. En 2019 el Gobierno federal contaba con unos 450 evaluadores cuyas competencias son evaluadas y certificadas por el Consejo del Tesoro.

Desde la década de 1970, el Código Presupuestario Federal dispone en Alemania que las auditorías deben valorar el logro de los objetivos, los impactos y la rentabilidad de las políticas públicas. Además, el Reglamento Interno Conjunto de los Ministerios, que regula la organización de los ministerios y la cooperación entre ellos y con los órganos constitucionales federales, exige la realización de la evaluación de impacto y el análisis coste-eficacia. Cada ministerio cuenta con una unidad especializada en evaluación, siendo éstas bastante activas, pero también se apoyan en entidades independientes para realizar sus análisis de evaluación. En el Parlamento, el Tribunal de Cuentas realiza una extensa labor de auditoría del gobierno federal. Cuenta con un equipo de ciencias sociales y del comportamiento con el apoyo directo de la Cancillería. La mayoría de las evaluaciones son realizadas por institutos de investigación no universitarios con vinculaciones a las estructuras gubernamentales y al mundo académico; algunos son altamente especializados sectorialmente.

En el Reino Unido, el Ministerio de Hacienda y el Cabinet Office desempeñan funciones de dirección informal de las evaluaciones (OCDE, 2021), si bien cada ministerio es responsable de planificar y ejecutar sus propias evaluaciones. Otros organismos públicos no vinculados a los ministerios también desempeñan tareas relacionadas con la evaluación. La Oficina Nacional de Auditoría es un organismo parlamentario independiente, que desarrolla una importante actividad de supervisión del gasto en todos los sectores de política pública. Especialmente destacable es el papel que desarrolla en materia de evaluación el Instituto de Estudios Fiscales, en concreto en el análisis presupuestario de los proyectos de ley de financiación o las propuestas de los principales partidos políticos en períodos preelectorales. En 2010 se creó Behavioural Insights Team con la finalidad de incorporar las ciencias del comportamiento en el diseño y evaluación de políticas públicas. Con frecuencia las evaluaciones se subcontratan con consultoras o centros de investigación buscando independencia. En 2004 se creó la Policy Profession, un grupo de empleados públicos con competencias certificadas en evaluación, que en 2017 contaba con casi 18.300 personas.

En Estados Unidos, desde 2018, la Oficina de Gestión y Presupuesto de la Casa Blanca cuenta con el Evidence Team, encargado de impulsar la evaluación y su coordinación dentro de la Administración federal, aunque los ministerios pueden decidir las evaluaciones que realizan y que incluyen en sus planes, para lo que cuentan con potentes unidades de análisis y evaluación. El Congreso de los Estados Unidos dispone de tres organismos que realizan actividades de evaluación: la General Accounting Office, el Servicio de Investigación, que realiza análisis legislativos, y una Oficina Presupuestaria. La Oficina de Ciencias de la Evaluación ha desarrollado un interesante material de apoyo,

tanto para la elaboración de las denominadas Agendas del Aprendizaje, como para el diseño de los planes de evaluación y el análisis de las capacidades de los ministerios. Las Agendas del Aprendizaje deben formar parte del plan estratégico de la agencia o ministerio e identificar qué temas de política pública necesitan evidencia y cómo se abordará su análisis, incluyendo la metodología que se utilizará. La Ley de Evidencia incorpora la Ley de Datos Abiertos del Estado, que obliga a los organismos a publicar información en línea usando formatos reutilizables.

En Francia, el Tribunal de Cuentas asiste al Parlamento y al Gobierno en la evaluación de políticas públicas. En 2009, se creó el Comité de Evaluación y Control para la coordinación de actividades de evaluación. Cada ministerio programa su propio plan de trabajo, aunque desde 2012, el Gobierno propone un programa de evaluación en el marco de la iniciativa de Modernización de la Acción Pública. Dos tercios de las evaluaciones se encargan a consultoras privadas, y unas pocas a las universidades, encargándose de las restantes distintos órganos de la administración. El Centro de Acceso Seguro a los Datos, 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.

