

Resumen

Este artículo discute algunas líneas metodológicas para la evaluación de la eficiencia y la productividad de los servicios públicos. Para ello, se repasa la utilidad de realizar *benchmarking* mediante métodos frontera y se presentan tres desarrollos metodológicos recientes; metafronteras y tecnologías de referencia, evaluaciones causales de los programas para la mejora de la eficiencia y la evaluación de las ganancias potenciales de realizar fusiones de unidades que prestan un mismo servicio. Se concluye que para iniciar esta tarea sería necesario crear agencias autonómicas independientes de evaluación para aquellos servicios en los que la comunidad autónoma tiene competencias.

Palabras clave: eficiencia, productividad, frontera de producción, servicios públicos.

Abstract

The aim of this paper is introducing some methodology research lines for evaluating the efficiency and productivity of public services. To do this, I review the advantages of performing benchmarking using frontier methods together with three recent methodological developments; meta-frontiers and the use of reference technologies, causal evaluations of public programs to improve efficiency and the potential gains from merging units inside an industry providing the same service. It is concluded that to start this evaluation task it would be necessary to create Independent Evaluation Agencies at Autonomous Community level for those services in which the Autonomous Community has transferred powers.

Keywords: efficiency, productivity, production frontier, public services.

JEL classification: D21, D24, H11, H75.

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

Daniel SANTÍN GONZÁLEZ

Universidad Complutense de Madrid

I. INTRODUCCIÓN

LA crisis generada por la pandemia de la COVID-19 vuelve a poner en el centro del debate económico y político a unas administraciones públicas (AA. PP.) que se enfrentan a mayores demandas de servicios públicos por parte de unos ciudadanos cada vez más exigentes y mejor informados. Asimismo, el escenario macroeconómico indica que España cerrará 2021 con una deuda y un déficit público en el entorno del 120 por 100 y por debajo del 8 por 100 del PIB, respectivamente. Todo ello obliga, tal como ordena el artículo 31.2 de la Constitución Española, a redoblar esfuerzos para gestionar los impuestos pagados por los contribuyentes respondiendo a los criterios de eficiencia y economía.

Desde este punto de vista legal, las leyes presupuestarias de la mayor parte de países desarrollados también recogen el mandato de eficiencia en la gestión de los recursos públicos. En el caso español, el artículo 26.1 de la Ley General Presupuestaria, correspondiente a los principios y reglas de programación presupuestaria, señala que: «La programación presupuestaria se regirá por los principios de estabilidad presupuestaria, sostenibilidad financiera, plurianualidad, transparencia, eficiencia en la asignación y utilización de los

recursos públicos, responsabilidad y lealtad institucional (...)».

Para cumplir con esta voluntad del legislador, que no es otra que la voluntad que emana de los ciudadanos, la evaluación de la actividad del sector público constituye un proceso complejo, pero absolutamente imprescindible, para cuantificar la eficiencia de las distintas actividades que realizan las AA. PP. y poder así optimizar los servicios que se prestan a los ciudadanos. Podemos definir la evaluación del sector público como el proceso sistemático de generación de información e indicadores relevantes para establecer, mediante el análisis económico, una relación entre recursos, actividades y resultados con el objetivo de determinar si los programas y/o servicios públicos ofrecidos a los ciudadanos resultaron eficaces, eficientes y socialmente rentables. Todo ello con el objetivo de que tanto los políticos como la sociedad civil puedan usar los resultados en su toma de decisiones.

Por evidente que parezca al hablar de evaluación es conveniente volver a subrayar que no existe una única aproximación estadística, econométrica o matemática para llevar a cabo la evaluación de todos los programas, políticas y servicios que desarrolla el sector público. Así, no será lo mismo evaluar si merece la pena extender la alta velocidad para llegar a una nueva

ciudad, evaluar un programa para mejorar la empleabilidad de los jóvenes que abandonaron la educación secundaria y reducir así su vulnerabilidad en el mercado de trabajo, o la evaluación de la eficiencia de un conjunto de centros de salud que operan en una determinada región.

En las dos últimas décadas ha surgido un creciente interés por los experimentos controlados aleatorizados y las técnicas de inferencia que emplean una situación contrafactual para el análisis y la evaluación de programas y políticas públicas. El objetivo de toda esta positiva línea de trabajo es mostrar si existe una relación causal entre la intervención del sector público, mediante una política o programa determinado, y el cambio en el indicador o indicadores asociados a la población beneficiaria del programa con relación a un grupo de control correctamente definido. Prueba de este interés y de la robustez de los resultados que están siendo obtenidos con estas técnicas son las recientes concesiones en 2019 (Barnejee, Duflo y Kremer) y en 2021 (Angrist, Card y Imbens) del Premio Nobel de Economía a profesores que desarrollan estas líneas de trabajo.

Siendo del todo indiscutible la importancia de esta dimensión de análisis causal en la evaluación de políticas públicas conviene también volver a recordar la importancia de llevar a cabo la evaluación sistemática de la eficiencia y la productividad de los servicios públicos mediante el uso de fronteras productivas. La oferta de servicios públicos que se realiza mediante unidades independientes de producción pública, que consumen trabajo y capital para entregar servicios de

manera satisfactoria a una población determinada, es ingente.

Redondeando, y únicamente a título ilustrativo, en España podemos contar unos 4.500 centros donde se imparte Educación Infantil, 9.800 centros de Educación Primaria, unos 4.000 centros de Educación Secundaria Obligatoria y Bachillerato, unos 460 hospitales, 3.000 centros de salud, 10.000 consultorios médicos locales, unos 1.000 juzgados de primera instancia e instrucción, 230 comisarías de policía, 130 servicios de bomberos, 68 cárceles, 47 universidades públicas presenciales, 46 puertos, 45 aeropuertos y un largo etcétera.

El gasto público asociado a la prestación de estos servicios es muy elevado y, además, a diferencia de «programas para la mejora de» que empiezan y en algún momento acaban, se repiten de forma sostenida año tras año. En 2018, el gasto presupuestado por las administraciones públicas para hacer frente a los servicios sanitarios (educativos) fue de unos 77.000 (45.000) millones de euros (Ministerio de Sanidad, 2021; MEFP, 2021). Esto supone que, solo en sanidad y educación, por cada punto porcentual que pudiéramos mejorar la eficiencia en la gestión de estos servicios liberaríamos 1.220 millones de euros bien para usos alternativos o bien en reducción de impuestos.

En el contexto de la evaluación del sector público, el interés por mejorar su eficiencia ha constituido uno de los principales focos de atención e investigación por parte de la economía pública, así como por investigadores de distintos ámbitos. Por motivos de extensión, se considera que no tiene sentido

volver a repasar ni los conceptos económicos de eficiencia y productividad ni todas y cada una de las metodologías de fronteras productivas (1) estándares disponibles para medir la eficiencia, tanto no paramétricas, incluyendo el análisis envolvente de datos (*data envelopment analysis-DEA*) como paramétricas, siendo el análisis con fronteras estocásticas (*stochastic frontier analysis-SFA*) la técnica más aplicada. Para este fin se propone al lector los trabajos previos de Álvarez (2001), Coelli *et al.* (2005), Fried *et al.* (2008) y Kumbhakar, Wang y Horncastle (2015) (2).

Tampoco tendría sentido utilizar unas pocas páginas para, desde una perspectiva sectorial, hacer un repaso pormenorizado de las distintas actividades que produce el sector público y que son susceptibles de ser analizadas con especial atención al caso español. Para ello, se refiere al lector al excelente número 95 de *Papeles de Economía Española* (Funcas, 2003) que llevó por título «Sector Público y Eficiencia» y al número 105 de la *Revista Economistas* (Colegio de Economistas de Madrid, 2005) titulado «La eficiencia de los servicios públicos: viejos problemas, nuevos enfoques».

El objetivo de este artículo es discutir algunas de las líneas metodológicas que considero que podrían ser utilizadas para la evaluación de la eficiencia y la productividad en el ámbito de la prestación de los servicios públicos. Para ello, en la segunda sección se repasa la utilidad de medir la eficiencia y la productividad en el sector público y cuál es la situación actual en España respecto a este tipo de evaluación. En la sección tercera se presentan tres desarrollos me-

todológicos e ideas que podrían ser utilizadas para llevar a cabo el *benchmarking* de unidades productivas. Esto incluye la cuantificación de las diferencias en productividad mediante metafronteras y tecnologías de referencia, las evaluaciones causales de los programas para la mejora de la eficiencia y la evaluación de las ganancias potenciales de realizar fusiones y escisiones de unidades de producción pública. Finalmente, en la última sección se ofrecen conclusiones y líneas futuras de trabajo.

II. LA MEDICIÓN DE LA EFICIENCIA Y LA PRODUCTIVIDAD DE LOS SERVICIOS PÚBLICOS EN ESPAÑA

Cualquier manual de microeconomía define una tecnología como la capacidad de obtener distintos *outputs* a partir del uso de múltiples *inputs*. Su caracterización básica se realiza mediante el concepto de función de producción. Así, una función de producción $y = f(x)$ indica la *máxima cantidad* y de *output* que es factible producir a partir de una determinada cantidad de *inputs* $(x_1, \dots, x_N) > 0$ (Mas-Colell, Whinston y Green, 1995, p.129). Se subraya «máxima cantidad» porque rara vez los manuales de microeconomía definen, asociado al concepto de función de producción, el concepto de eficiencia.

En efecto, en mercados que operan en competencia perfecta y con productores privados precio-aceptantes guiados por la maximización del beneficio, no caben resultados ineficientes ya que estos llevarían directamente a la quiebra a estas empresas y, por tanto, se asume que la pro-

ducción se lleva a cabo con plena eficiencia. En principio, nada impediría que la producción de bienes y/o servicios públicos alcance idénticos resultados. Sin embargo, por distintas razones asociadas a la gestión de las unidades productivas las unidades de producción públicas encuentran dificultades para alcanzar los máximos teóricos permitidos por la tecnología (Farrell, 1957; Leibenstein, 1966; Levin, 1974; Niskanen, 1975; Levin, Jamison y Radner, 1976). La multiplicidad de objetivos heterogéneos y a veces no definidos, la falta de restricciones presupuestarias claras y la estructura de incentivos pueden apartar a la empresa pública de alcanzar la máxima eficiencia (Tirole, 1994).

Ello implica que es factible que las unidades pueden llevar a cabo la producción de los servicios públicos de manera ineficiente, esto es, utilizando más recursos de los que se requieren para lograr una cantidad de producción o, alternativamente, dependiendo de la orientación elegida, produciendo por debajo del máximo tecnológicamente factible dados unos recursos. A partir de este hecho, y dado que rara vez se observan quiebras en los sectores enumerados en la sección anterior, surge la necesidad de conocer la magnitud de la eficiencia con la que se lleva a cabo la producción de un servicio público comparando a las distintas unidades productivas que lo desarrollan.

La utilidad de esta evaluación vista la magnitud presupuestaria que supone la prestación de estos servicios, es considerable. En este sentido, el australiano Steering Committee for the Review of Government Service Provision (SCRGSP, 2013) señala que tra-

dicionalmente todos los esfuerzos de las administraciones públicas para mejorar los servicios se han centrado en incrementar los recursos presupuestarios, como si fuera evidente que más euros invertidos automáticamente fueran capaces de producir más y mejores servicios. Sin embargo, el objetivo de las AA. PP. debería cambiar y buscar las mejores maneras de utilizar los recursos públicos existentes (muchos o pocos, su magnitud, dependerá de las preferencias democráticas que desean los ciudadanos acerca del peso del sector público que quieren para la economía donde residen) desplazando el foco de atención desde la cuantía de los recursos, «más» no supone tautológicamente «mejor», a un uso de los mismos eficiente.

Las principales ventajas de llevar a cabo la evaluación del desempeño de los servicios públicos mediante competencia por comparación son numerosas, pero pueden destacarse las siguientes. En primer lugar, ayuda a definir el cumplimiento de objetivos, tanto cuantitativos como cualitativos, así como a clarificar las responsabilidades de las organizaciones. En segundo lugar, anima a promover el análisis de las relaciones entre agencias y programas para una mejor coordinación de las políticas dentro de las AA. PP. También mejora la transparencia, ofreciendo tanto a los distintos Gobiernos, central, autonómico y local, como a los ciudadanos indicadores de desempeño de sus políticas y servicios públicos a lo largo del tiempo. En tercer lugar, la competencia por comparación anima a la continua mejora en el desempeño ya que, conocidos los indicadores de desempeño, ninguna unidad productiva querrá

aparecer la última en el *ranking*. De esta manera el *benchmarking* permite identificar a las mejores prácticas, aprender de su gestión y promover la introducción de mejoras e innovación en la entrega de los servicios. Y no menos importante, también permite identificar a las peores prácticas, con el objetivo de poner remedio a los problemas que impiden que ofrezcan el servicio de manera eficiente, mediante el establecimiento de objetivos de mejora.

Hasta la fecha, la institucionalización de la evaluación sistemática de los servicios públicos en España ha tenido escaso éxito. El primer intento serio de llevar a cabo una evaluación de la eficiencia del gasto público tuvo su origen en la creación, a mediados de 2006, de la Agencia Estatal de Evaluación de las Políticas Públicas y la Calidad de los Servicios (AEVAL). La misión de AEVAL era «la promoción y realización de evaluaciones y análisis de impacto de las políticas y programas públicos, así como el impulso de la gestión de la calidad de los servicios, favoreciendo el uso racional de los recursos y la rendición de cuentas a la ciudadanía». Sin embargo, las evaluaciones planteadas por AEVAL no incluían los servicios prestados por las comunidades autónomas (CC. AA.), solo los de la Administración General del Estado, y no respondieron claramente a los objetivos definidos, por lo que la Agencia fue definitivamente clausurada en 2017.

En 2016, como consecuencia de la crisis financiera global, España adquirió el compromiso ante la Comisión Europea de revisar la eficiencia de las administraciones públicas. Esta tarea de análisis y revisión del gasto público conocida como *spending review*

es encargada por el Gobierno a la Autoridad Independiente de Responsabilidad Fiscal (AIReF) que había sido creada en 2014. El primer Plan de Actuación del *spending review* desarrollado por AIReF marcó un punto de inflexión al incluir, por primera vez de manera explícita en un texto de una institución pública, la posibilidad de realizar *benchmarking* mediante el uso de fronteras productivas, lo que AIReF (2017, p. 6) denomina «frontera eficiente». Concretamente, el Plan de Acción de AIReF afirma: «Todo ello con objeto de medir tanto el “coste medio” de la intervención pública como “gasto medio” por beneficiario. Estas evaluaciones también pueden incluir el cálculo de fronteras eficientes de producción, que determinan un *benchmarking*». Si bien estos análisis han empezado a aparecer en algunas de sus evaluaciones, como en AIReF (2018) para el análisis de la Sociedad Estatal Correos y Telégrafos o en AIReF (2020) para el estudio del sistema universitario público andaluz, se trata todavía de análisis muy incipientes que distan de ser sistemáticos (3). En la siguiente sección se apuntan algunas ideas para institucionalizar el *benchmarking* en las AA. PP.

III. LA EVALUACIÓN DE LA EFICIENCIA Y LA PRODUCTIVIDAD EN EL SECTOR PÚBLICO

Una vez realizadas las primeras aproximaciones con algunas evaluaciones puntuales, el siguiente paso es institucionalizar la evaluación del desempeño de la producción de los servicios públicos. Es decir, la evaluación de la eficacia, la eficiencia y la productividad de manera continuada en el tiempo, y que en

cada ejercicio explora la mejora de los mecanismos de búsqueda y análisis de los indicadores asociados a esta evaluación. Para ello, el primer requisito siempre debe ser la definición de los *inputs* y *outputs* que serán utilizados y que idealmente deberían incluir indicadores de cantidad, calidad (4), *outputs* no deseables (5) así como variables que tengan en cuenta el entorno en el que operan (6) las distintas unidades. En la mayoría de los casos la literatura es amplia y las posibilidades son múltiples, por lo que se recomienda alcanzar el mayor consenso posible entre evaluador, unidades evaluadas y usuarios finales de la evaluación teniendo en cuenta el contexto y los objetivos perseguidos. Ello también incluye la realización de distintos análisis para comprobar la robustez en los resultados.

A continuación, deben definirse las fuentes de datos de donde se obtendrá la información necesaria para ejecutar los análisis. En numerosas ocasiones las evaluaciones de eficiencia y productividad no se llevan a cabo por parte de las AA. PP. o por parte de otros agentes argumentando que no existe información acerca de todas las variables con las que sería necesario contar. En este sentido, es importante señalar que siempre existirá información e indicadores no disponibles y también siempre será posible perfeccionar las medidas ya existentes. Este hecho genera un círculo vicioso ya que los indicadores no se generan porque no existe evaluación y la evaluación nunca empieza porque se considera que no existen los indicadores adecuados. El resumen es que para romper este círculo es imprescindible la voluntad política de comenzar a evaluar.

Como ya se ha señalado, la evaluación es un proceso de generación de información relevante continuo en el tiempo y, por tanto, posiblemente la información para el análisis nunca será totalmente perfecta. Comenzar a evaluar supone diseñar sistemas de información para almacenar y gestionar los datos necesarios en la evaluación. Inicialmente estos serán imperfectos, pero con el paso del tiempo las medidas serán cada vez más precisas y adecuadas para conseguir mejores evaluaciones. Un ejemplo de la puesta en servicio de un sistema de información sería euroCRIS, (Current Research Information Systems for Europe), que recoge datos sobre la investigación llevada a cabo por distintas instituciones y universidades a partir de las actividades de sus profesores e investigadores.

En este ejemplo, una vez obtenida esta información, la misma puede ser utilizada para analizar la eficiencia y la productividad del sistema de investigación español en comparación con otros países europeos. Además, el sistema permitiría también comparar los centros de investigación en España. Por supuesto, la puesta en marcha de un sistema de información llevará tiempo, pero esta no parece una dificultad insalvable en la sociedad interconectada del siglo XXI.

Obtenida la información cabe plantearse cuáles son los objetivos de análisis que se pueden desarrollar. A continuación, señalo tres líneas de trabajo que, en mi opinión, serían especialmente deseables para mejorar la toma de decisiones: 1) las comparaciones y evolución de la productividad total de los factores; 2) la evaluación causal de programas

públicos para la mejora de la eficiencia y la productividad; y 3) el análisis de la fusión y escisión de unidades que prestan servicios públicos.

1. Metafronteras y diferencias en productividad a lo largo del tiempo

Como ya se ha señalado, la medición de la eficiencia con la que operan los servicios públicos es una parte fundamental para proporcionar los mejores servicios posibles a los ciudadanos haciendo un buen uso de los recursos. Además de los objetivos ya señalados, Pedraja-Chaparro, Salinas-Jiménez y Smith (1999) señalan como objetivo de esta evaluación «el análisis de la eficiencia global de la industria». En el caso del sector público español entendemos por industria al conjunto de unidades productivas que producen un servicio en un determinado sector que está bajo la jurisdicción de una misma administración (7). Algunos ejemplos podrían ser los institutos de Educación Secundaria (IES) de la Comunidad de Castilla y León, las oficinas de la empresa pública Correos o los hospitales andaluces. Estos centros forman «una industria» que depende presupuestariamente de una Administración interesada en ofrecer el mejor y más eficiente servicio a sus ciudadanos.

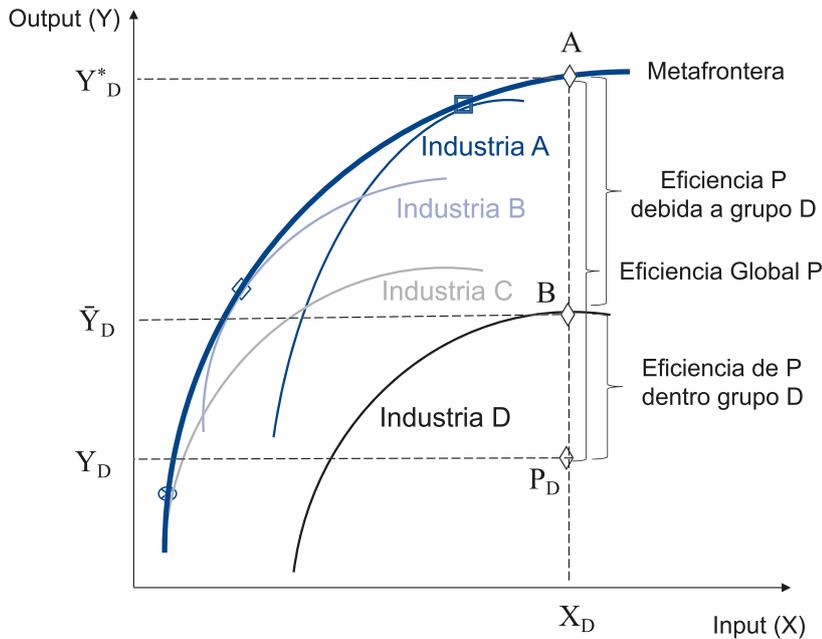
A partir de aquí el primer objetivo de estas administraciones es analizar la eficiencia de «su industria» y cómo evoluciona a lo largo del tiempo. A continuación, podemos comparar el desempeño de la industria con otras industrias semejantes. Por ejemplo, podemos analizar las diferencias en desempeño de los

IES con los de los centros de educación concertada de la misma CC. AA. que prestan con dinero público los mismos servicios. También podemos comparar el desempeño de grupos de centros que operan en municipios de distinto tamaño o incluso comparar el desempeño de médicos dentro de un mismo hospital.

Siendo estas primeras mediciones de la eficiencia y de los cambios en productividad a lo largo del tiempo importantes, el siguiente paso sería realizar el análisis de manera comparada no solo dentro de cada industria, sino entre industrias que pertenecen a distintas jurisdicciones. Que parezca que nuestra industria va bien será una medida vacía si nunca comparamos sus resultados con los que obtienen otras industrias de nuestro entorno (8). Podríamos entonces comparar industrias de distintas comunidades autónomas o incluso de distintos países a lo largo del tiempo.

Por tanto, se trata de recuperar la idea de «federalismo de laboratorio» como una importante ventaja del sistema autonómico español todavía no bien explotada. Para Oates (1999), un sistema descentralizado permite probar y experimentar distintas alternativas de prestar un servicio público. Esta experimentación tendría por objetivo obtener ganancias de eficiencia en términos de progreso tecnológico que a su vez animen a más experimentación. En definitiva, las CC. AA. podrían introducir distintas innovaciones en sus servicios para maximizar el nivel de los *outputs* que entregan a sus ciudadanos dado un nivel de gasto, al tiempo que reproducen los programas exitosos en otras CC. AA. y descartan aquellos cambios en

GRÁFICO 1
EL CONCEPTO DE METAFRONTERA



Fuente: Elaboración propia.

los programas que no suponen ganancias de eficiencia.

Para realizar este «federalismo de laboratorio» acerca de lo que funciona y lo que no funciona en términos de mejorar la eficiencia y la productividad de los servicios podemos recurrir al uso de metafronteras (para una revisión se recomienda las lecturas de Battese y Prasada-Rao, 2002; Pastor y Lovell, 2005 y Oh, 2010).

El concepto de metafrontera recoge la idea de trazar una única frontera a partir de las mejores prácticas de las unidades productivas pertenecientes a las distintas industrias y podemos ilustrarlo a partir del gráfico 1.

En el gráfico 1 se representan las fronteras de producción de cuatro industrias; A, B, C y D, capaces de producir un *output* a partir de un *input*, cada una de

ellas formada por un número de unidades de producción. La idea de metafrontera implica trazar una nueva frontera productiva asumiendo que la combinación lineal de las mejores unidades productivas pertenecientes a los distintos grupos o industrias es tecnológicamente factible. Así, la eficiencia de cualquier unidad productiva se podría descomponer en dos partes, una parte debida a su propia gestión en comparación con otras unidades productivas que compiten con ella en la misma industria y una parte debida a la diferencia tecnológica entre las industrias comparadas.

La unidad productiva no sería directamente responsable de esta segunda parte de la distancia a la frontera, sino que lo serían los técnicos y políticos encargados de la industria, que deberían analizar el porqué de

estas diferencias. Por un lado, las causas podrían ser exógenas, dependientes de variables geográficas, de dispersión de la población o en cualquier caso de factores fuera del alcance de los gestores de la industria. En este caso, existen distintas metodologías para incorporar directamente estas variables en el análisis. Finalmente, la parte no explicada por estos factores exógenos sería eficiencia en la gestión o cuellos de botella debidos a marcos regulatorios que impiden el pleno desarrollo de la industria para favorecer otros objetivos distintos de la eficiencia.

Por ejemplo, en el gráfico 1 P_D sería una unidad ineficiente ya que la frontera productiva D construida para su industria indica que con la cantidad de *input* que utiliza podría aumentar la cantidad actual que produce de *output* Y_D hasta el nivel \bar{Y}_D . Además, si comparamos a la industria D con la metafrontera, construida para los restos de unidades analizadas, observamos cómo existe un objetivo de producción potencial que se sitúa en Y^*_D . Siendo la diferencia entre Y^*_D y \bar{Y}_D debida a la industria. Las metafronteras han sido empleadas en España para el análisis de servicios locales de recogida de basura (Garrido-Rodríguez *et al.*, 2018), el transporte público (Campos-Alba, 2020) o para evaluar diferentes sistemas educativos (Cordero, Santín y Simancas, 2017; Giménez *et al.*, 2019), entre otros.

La metafrontera puede ser estimada mediante distintos procedimientos como una frontera estocástica, DEA u orden-m, pero un supuesto básico para el uso de esta metodología es asumir que la tecnología sintética construida en la metafrontera

es factible y reproducible en la industria que se desea mejorar. Esto en la práctica puede suponer problemas prácticos ya que las unidades productivas en la metafrontera no operan en los mismos grupos y por tanto la referencia es difícil de copiar ya que sus *DMUs* (*decision making units* o unidades de decisión) operan en distintas industrias bajo distintos marcos legales, contextos sociales o económicos, lo que puede llegar a hacer imposible alcanzar la referencia.

Para superar esta dificultad en la comparación de industrias, Camanho y Dyson (2006) proponen adaptar el índice de Malmquist (Caves, Christensen y Diewert, 1982) donde los dos períodos de tiempo que se utilizan en la definición de este índice se sustituyen por dos grupos de unidades productivas que son comparados. El objetivo de estos autores es superar esta dificultad de las metafronteras para hacer comparaciones directas de las diferencias en productividad entre

industrias. El gráfico 2 muestra un ejemplo para entender mejor esta metodología.

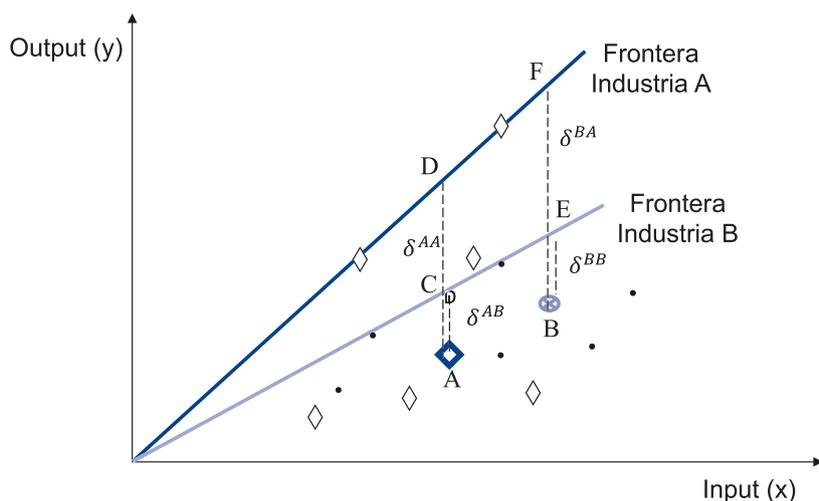
En el gráfico 2 observamos las fronteras productivas de dos industrias A y B cuyas unidades están representadas por rombos y círculos respectivamente. Por un lado, podemos ver cómo la frontera productiva de la industria A domina en términos tecnológicos a la de la industria B, ya que con la misma cantidad de *input* es capaz de obtener mayor *output*. Para cada unidad productiva podemos calcular este índice de Malmquist adaptado, como la diferencia en productividad entre dos industrias, A y B, en vez de entre dos períodos, t y $t + 1$ como haríamos normalmente con aquel índice.

En el gráfico 2 los puntos A y B representan la posición hipotética promedio de las dos industrias, determinadas a partir de las medias geométricas de las distancias de las unidades de cada industria respecto a las dos fronteras. Esto

nos permite descomponer las diferencias en productividad entre las dos industrias, como en el índice de Malmquist, en una parte debida a diferencias tecnológicas y otra a diferencias en eficiencia. Así, podemos comprobar cómo, aunque por causas tecnológicas (media geométrica de las distancias DC y EF) la productividad en A es mayor que en B, la productividad de la industria B es mayor que la A en términos de eficiencia (la distancia BE es menor que la distancia AD y, por tanto, está más cerca de la frontera). El producto de la diferencia en productividad debida a tecnología y a eficiencia nos dará una medida acerca de cuál es la diferencia global en productividad entre las industrias. Esta información es muy útil para aprender qué podemos hacer para mejorar la productividad del servicio que ofrecemos a los ciudadanos a partir de la comparación con otras industrias que también lo desarrollan.

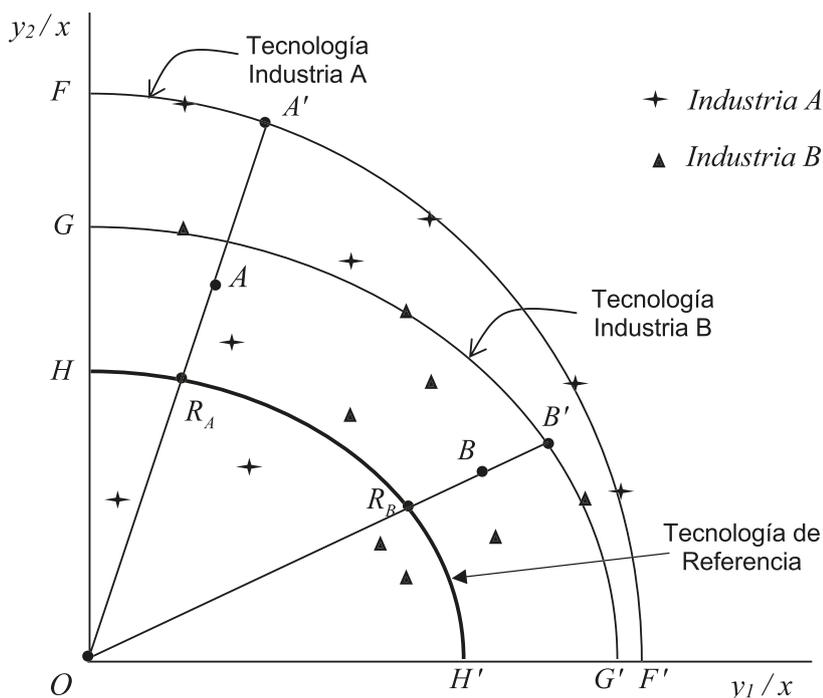
El índice de Camanho y Dyson es atractivo y ha sido utilizado para evaluar hospitales según la titularidad con la que actúan (Ferreira y Cunha-Marques, 2015) o la gestión de aeropuertos (Ferreira, Marques y Pedro, 2016). Sin embargo, el índice presenta dos inconvenientes para su uso sistemático en la evaluación del sector público. En primer lugar, no se pensó para hacer seguimientos de las diferencias a lo largo del tiempo. En segundo lugar, con más de dos grupos se rompe la condición deseable de circularidad de Frisch (1936) en números índices. Es decir, dados tres grupos A, B y C, la diferencia global en productividad, y en sus componentes entre los grupos A y C, debe ser el producto de las diferencias entre A y B por una parte y B y C por otra.

GRÁFICO 2
DIFERENCIAS EN PRODUCTIVIDAD ENTRE INDUSTRIAS
(CAMANHO Y DYSON, 2006)



Fuente: Elaboración propia.

GRÁFICO 3
DIFERENCIAS EN PRODUCTIVIDAD ENTRE INDUSTRIAS
(APARICIO Y SANTÍN, 2018)



Fuente: Elaboración propia.

Para solucionar este problema, Aparicio y Santín (2018) proponen un nuevo índice basado en Berg, Førsund y Jansen (1992) que consiste en introducir la tecnología de un «grupo de referencia» (R). Con ello, los problemas de circularidad entre dos o más grupos a lo largo del tiempo desaparecen. Otra ventaja de este índice es que no es necesario actualizar la metafrontera en cada nuevo período, lo que permite hacer un seguimiento sistemático y robusto de las diferencias de productividad entre industrias a lo largo del tiempo (9). El gráfico 3 muestra las ideas básicas y las ventajas de introducir una referencia para el cálculo de este índice.

En el gráfico 3 encontramos nuevamente dos industrias A y B que utilizan un único input

para producir dos *outputs*. Al igual que en los gráficos 1 y 2, el gráfico 3 muestra las fronteras productivas definidas por las tecnologías de cada industria («FF» y «GG» para las industrias A y B respectivamente). Pero, para realizar las comparaciones, en lugar de utilizar una metafrontera utilizamos la tecnología (HH') de un grupo de unidades productivas que las industrias que serán evaluadas consideran una referencia. La proyección de las medias geométricas de las unidades de cada sector, representadas en el gráfico 3 por los puntos A y B, respecto a su propia tecnología (AA» y BB') y respecto a la tecnología de referencia (AR_A y BR_B) permiten medir cómo las diferencias en productividad de las dos industrias evolucionan a lo largo del tiempo y explicarlas

debido a los cambios tecnológicos y de eficiencia que se irán produciendo.

Independientemente de la metodología matemática o econométrica elegida para la estimación final, el seguimiento sistemático de unidades productivas y el análisis de las diferencias en la productividad de las industrias a lo largo del tiempo debería ser un objetivo prioritario para institucionalizar la medición del desempeño de los servicios públicos.

2. Hacia una evaluación causal de los cambios en eficiencia y productividad

El siguiente paso, una vez lograda la institucionalización de la medición de los cambios en productividad de los distintos servicios públicos, será el análisis causal acerca de qué políticas públicas permiten mejorar el desempeño de las unidades productivas. Para ello, una línea actual de investigación consiste en combinar las herramientas tradicionales de inferencia causal con las técnicas de medición de la eficiencia basadas en el uso de fronteras. En este sentido, Mergoni y De Witte (2022) revisan 81 trabajos que realizan análisis causales en este sentido. De ellos, 53 trabajos tienen por objetivo analizar si cambios en políticas públicas o la aplicación de nuevos programas afectaron al desempeño de las unidades de producción respecto a un grupo de control. La mayoría de las aplicaciones reportadas son en el ámbito educativo [22], seguidas de servicios locales y ayuntamientos [13] y del sector salud [10], siendo el método de diferencias en diferencias el más aplicado (10).

Además de la combinación de fronteras productivas y técnicas de inferencia causal, una vez institucionalizada la medición del desempeño, en el futuro, también podrían diseñarse experimentos controlados aleatorizados para evaluar mejoras en productividad. En este sentido, Santín y Sicilia (2017) señalan cómo un nuevo programa público puede tardar bastante tiempo, debido a la necesidad de ganar experiencia o a problemas de gestión y ejecución que surgen en la aplicación de procedimientos novedosos, en producir diferencias en resultados estadísticamente significativos en la media en los *outcomes* deseados.

Sin embargo, estos autores confirman mediante un experimento de Monte-Carlo que es más fácil evaluar en el corto plazo si el programa está consiguiendo mejoras en productividad debido a causas tecnológicas, esto es, porque las mejores prácticas del grupo tratado empiezan a obtener mejores resultados que las mejores prácticas del grupo de control, o de eficiencia, cuando la eficiencia media de las unidades tratadas es significativa-

mente mejor que las del grupo de control. El gráfico 4 resume estas ideas.

El panel A del gráfico 4 muestra las unidades productivas de una industria que utiliza una única tecnología para producir una cantidad de *output* a partir de una cantidad de *input*. Mediante un proceso de aleatorización podemos dividir a las unidades en dos grupos; de tratamiento (T) y de control (C), representados en el gráfico por círculos blancos y negros respectivamente, cuyas medias en *inputs*, *outputs* y niveles de eficiencia no son estadísticamente distintas en el período base. El panel B del gráfico 4 muestra que si bien después de la aplicación del programa la eficiencia media de ambos grupos no es estadísticamente distinta $\bar{u}_T \cong \bar{u}_C$, la tecnología del grupo tratado domina ahora a la del grupo de control, indicando que el programa, si es bien implementado, puede mejorar la productividad de los servicios públicos. A partir de este análisis, y mediante una evaluación más detallada y cualitativa, se puede analizar las razones y canales que han motivado que estas unidades hayan mejo-

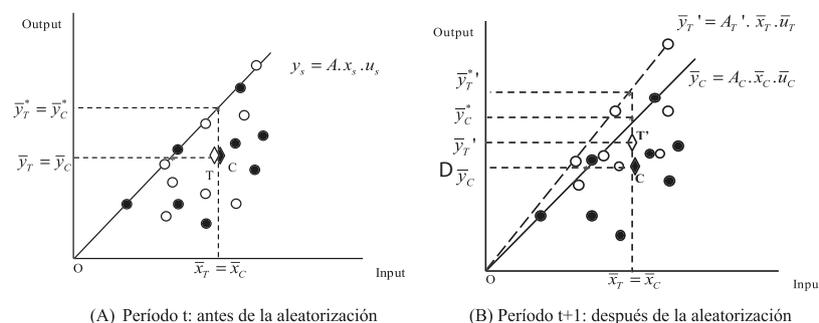
rado, aprender de su experiencia y trasladar estas buenas prácticas al resto del grupo tratado en períodos sucesivos.

3. Fusiones y escisiones de unidades productivas

Una decisión relevante en una institución es la escala con la que llevará a cabo sus operaciones. Pensemos, por ejemplo, en el caso de una universidad pública. Una universidad normalmente no opera de manera centralizada, lo habitual es que opere a partir de distintas facultades que, por ejemplo, agrupan distintas titulaciones relacionadas entre sí. Además, los profesores con intereses investigadores y/o docentes que pertenecen a una misma área o ámbito de conocimiento suelen agruparse por departamentos que a su vez pueden pertenecer a una o varias facultades. Pero, ¿cuál es el número óptimo de facultades y departamentos en una universidad (o en cualquier otra entidad pública que se organizara de una manera similar)?

La respuesta a esta pregunta es compleja ya que cada unidad de servicio supone distintos costes fijos y variables asociados a su proceso productivo. En este caso, al fusionar departamentos se podrían reducir los costes variables al liberar puestos de personal de administración y servicios de los departamentos fusionados para que realicen otras tareas o se pueden obtener ahorros por la reducción de complementos retributivos o de horas liberadas en gestión que ahora se podrían dedicar a tareas docentes e investigadoras. Por otro lado, tampoco parece sensato centralizar masivamente todos los servicios en un único departamento. Los costes de

GRÁFICO 4
ESQUEMA DE UN EXPERIMENTO CONTROLADO ALEATORIZADO UTILIZANDO FRONTERAS DE PRODUCCIÓN



Fuente: Elaboración propia a partir de Santín y Sicilia (2017).

tiempo en coordinación podrían ser tan elevados que merecería la pena dividir la unidad en varias unidades que operen de manera independiente. La casuística es elevada y, por tanto, es necesario sistematizar la búsqueda de las ganancias o pérdidas potenciales para cada caso.

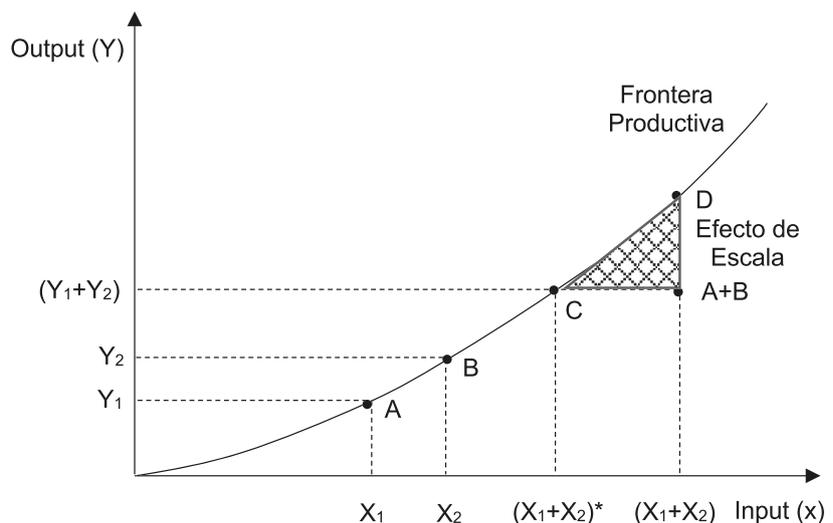
Las técnicas de fronteras productivas pueden ser utilizadas para analizar estas ganancias esperadas en eficiencia como consecuencia de fusionar o dividir unidades que gestionan el mismo servicio. En principio, en este ejemplo, una fusión entre varios departamentos será deseable si el nuevo departamento es capaz de lograr obtener una mayor producción en términos de horas de docencia y resultados de investigación o de alcanzar la misma producción a un menor coste respecto a los departamentos iniciales.

Siguiendo a Bogetoft y Wang (2006) y a Bogetoft y Otto (2011), las ganancias totales de eficiencia asociadas a la fusión pueden ser descompuestas en tres ganancias potenciales; efecto aprendizaje, efecto de armonización y efecto tamaño. El primer efecto es el denominado efecto aprendizaje (*learning effect*), o también llamado efecto eficiencia, y muestra cuánto puede mejorar la eficiencia del servicio si la nueva unidad opera en la frontera productiva. Obviamente, si las dos unidades son fusionadas «por decreto-ley», pero nada cambia, seguramente esta ganancia potencial será equivalente a la ganancia que las unidades previas alcanzarían si también alcanzaran la eficiencia. Sin embargo, la fusión podría buscar cambiar inercias negativas en la gestión que no cambiarían si no se introducen cambios (11).

La segunda ganancia potencial es la denominada efecto escala o efecto tamaño (*scale effect*). Este efecto se ilustra en el gráfico 5 donde se representa una tecnología en la que existen dos unidades productivas A y B que operan de manera plenamente eficiente al estar situadas sobre la frontera productiva. Si existen economías de escala, y por tanto con un tamaño mayor se pueden reducir los costes medios, la fusión de A+B permitiría llevar a cabo la misma cantidad de producción (Y_1+Y_2), pero reduciendo los inputs utilizados desde (X_1+X_2) hasta $(X_1+X_2)^*$, siendo $(X_1+X_2)^* < (X_1+X_2)$. En el gráfico 5 la ganancia potencial vendría dada por el triángulo (A+B)CD. Cabe destacar que esta ganancia podría ser negativa. El resultado final dependerá de la tecnología productiva que determinan las unidades productivas antes de la fusión, tecnología que deberá ser analizada mediante la construcción de una frontera productiva. De ahí la importancia de la evaluación.

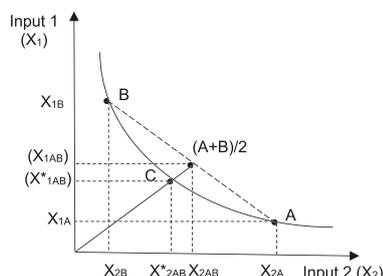
La tercera ganancia potencial es el efecto de armonización (*harmony effect*) que podemos ilustrar en el gráfico 6. En ella, se representa una isocuanta unitaria que reproduce la combinación de dos *inputs* capaces de producir una unidad de *output*. Se puede observar como las unidades A y B están muy desequilibradas en cuanto a la intensidad del uso de los factores productivos ya que A (B) utiliza muchas (pocas) unidades del *input* 1 y pocas (muchas) del *input* 2. En este caso, la fusión de ambas unidades genera una nueva unidad productiva, representada por el punto $(A+B)/2$. Tras la fusión vemos como la tecnología permite una ganancia de eficiencia potencial ya que la unidad fusionada podría reducir equiproportionalmente sus *inputs* hasta el punto C liberando estos recursos para usos alternativos y continuar produciendo la misma cantidad de producción.

GRÁFICO 5
EFECTO TAMAÑO O DE ESCALA EN LA FUSIÓN DE UNIDADES DE GESTIÓN



Fuente: Adaptado de Bogetoft y Otto (2011, p. 266)

GRÁFICO 6
EFECTO ARMONIZACIÓN
EN LA FUSIÓN DE UNIDADES
DE GESTIÓN



Fuente: Adaptado de Bogetoft y Otto (2011, p. 270).

Las aplicaciones potenciales de esta metodología son numerosas tanto dentro de las organizaciones o servicios como entre ellos. Un ejemplo sería el de la fusión de municipios. España cuenta con más de 8.100 municipios y de ellos aproximadamente el 60 por 100 tiene menos de 1.000 habitantes. Salvo alguna excepción, como la fusión entre Oza dos Ríos y Cesuras, dos municipios de la provincia de A Coruña llevada a cabo en 2013, y la lograda entre las localidades de Don Benito y Villanueva de la Serena de la provincia de Badajoz en 2022, existen pocas iniciativas para la fusión de ayuntamientos.

Por ejemplo, para el caso de la integración de estos dos municipios pacenses, Cordero Pedraja-Chaparro y Suárez Pandiello (2022) señalan, entre otras numerosas ventajas, las ganancias en economías de escala; que incluye una posición más favorable para negociar con proveedores, reducción de costes medios y ganancias de financiación por aumento de población, al incrementarse del 1,3 al 1,5 el coeficiente multiplicador aplicado a los ingresos procedentes de

la participación en los ingresos tributarios del estado. Ganancias en eficiencia, al reducirse los costes en administración y mejorar la gestión de los recursos técnicos y humanos. Por ejemplo, la Ley de Régimen Electoral General concede 21 concejales para dos municipios entre 20.001 y 50.000 habitantes, de ellos entre 10 y 11 podrían tener dedicación exclusiva y percibir remuneración. Una eventual fusión por encima de 25.000 habitantes aumentaría a 25 el número de concejales y a 15 aquellos con dedicación a tiempo completo, lo cual supone un ahorro de recursos respecto a la suma de los dos municipios iniciales. Finalmente, la fusión permite ganancias de armonización, a la hora de alcanzar una mayor coherencia en la planificación económica del municipio aprovechando los recursos humanos e infraestructuras ya existentes, en lugar de proceder de manera individual a mejorar las carencias que tuviera cada municipio.

En definitiva, el análisis comparado de desempeño en el sector público permitiría cuantificar las ganancias y costes que se podrían obtener de potenciales fusiones para que una vez calculadas los ciudadanos y políticos decidan con mejor información si desean continuar adelante con la propuesta de fusión.

IV. CONCLUSIONES

La discusión realizada en las secciones precedentes pretende poner de manifiesto que:

– Siendo sumamente importante la evaluación de programas mediante inferencia causal basada en el uso de contrafactuales, la evaluación sistemática de la eficiencia y la productividad de

las unidades productoras de servicios públicos generaría fuertes ganancias de información para la toma de decisiones. Como ya se ha apuntado, un punto porcentual de mejora en la eficiencia de los servicios sanitarios y educativos supondría liberar 1.220 millones de euros para usos alternativos. Aunque en los últimos años se han dado pasos tímidos en el buen sentido, España está todavía lejos de alcanzar este objetivo.

– El primer requisito, en mi opinión, para llevar a cabo una adecuada evaluación sería crear agencias autonómicas independientes de evaluación de todos los programas y servicios públicos que dependan de los parlamentos autonómicos. En España, las CC. AA. tienen las competencias en sanidad y educación, los dos principales motores de gasto público «en especie». Por tanto, la administración autonómica, que es quien asigna el presupuesto, debería ser también la encargada en evaluar «sus propias industrias» (12). El objetivo inicial de cada agencia, que podría agrupar simultáneamente a varias CC. AA. si así lo deciden, sería disponer de información de calidad de manera estable en el tiempo para poder construir los indicadores necesarios para los análisis. Estos indicadores serían transparentes y estarían disponibles para que terceras partes pudieran realizar sus propios análisis. Este sistema de evaluación autonómica puede generar en el futuro problemas de coordinación e intercambio de información con agencias de otras CC. AA., pero este será un problema a medio plazo y de más fácil solución, al ya existir los mecanismos que generan la información.

– Construido el sistema anterior, el siguiente paso natural sería que varias CC. AA. se coordinen para comparar sus servicios utilizando los mismos indicadores. Ello pasará por una definición fina y común de los indicadores de consenso elegidos que permita la comparación. Con el tiempo, no entrar en el sistema interterritorial de evaluación tendrá un coste de reputación, como por ejemplo lo tendría salir de las evaluaciones internacionales PISA, TIMSS o PIRLS.

– Finalmente, levantado el sistema de información para cada servicio, las CC. AA. podrán explorar las ganancias (o pérdidas) en productividad que sus nuevas políticas tienen o tendrían en los servicios. Para ello se utilizarían, entre otras, las aproximaciones comentadas en la sección tercera. El indicador final de eficiencia junto a otras dimensiones, por ejemplo, de eficacia (mirando únicamente el lado de los *outputs*, de calidad o de reducción de *bad outputs*), desigualdad en la recepción de los servicios, etc., ofrecerá mayor transparencia para explicar las decisiones que afectan al servicio y al final los ciudadanos decidirán con sus votos la orientación final de las políticas.

Para terminar, no está mal resumir la importancia de evaluar utilizando la frase atribuida al célebre físico y matemático William Thomson Kelvin quien en el S. XIX afirmó: «Lo que no se define no se puede medir. Lo que no se mide, no se puede mejorar. Lo que no se mejora, se degrada siempre».

NOTAS

(1) A lo largo de este artículo, conceptos en economía como eficiencia, productividad, eficacia o evaluación del desempeño, se

entenderán en general referidos en relación con una frontera productiva o tecnología que subyace al proceso productivo y que debe ser estimada para permitir el *benchmarking* entre las unidades productivas comparadas.

(2) Abordar el detalle de todas y cada una de las últimas novedades metodológicas paramétricas y no paramétricas para medir eficiencia y productividad excede los objetivos de este artículo. Para comenzar con esta laboriosa tarea recomiendo el manual de SICKLES y ZELENYUK (2019).

(3) Para una excelente revisión del panorama actual y del marco regulador de la evaluación de políticas públicas en España se puede consultar DE LA FUENTE *et al.* (2021).

(4) No es el objetivo de este artículo realizar una revisión pormenorizada de todos los *inputs* y *outputs* que pueden ser utilizados en cada sector y de las metodologías que principalmente se usan para su evaluación. Para una revisión de variables y metodologías aplicadas en los principales sectores donde opera el sector público puede consultarse DE WITTE y LÓPEZ-TORRES (2017) para educación; HOLLINGSWORTH (2008) para el sector salud; NARBÓN-PERPIÑÁ y DE WITTE (2018a y 2018b) para servicios locales y ayuntamientos, y ÉSPASA y ESTELLÉ-MORÉ (2015) para juzgados de primera instancia.

(5) Además de los clásicos *inputs* y *outputs*, el análisis de eficiencia permite la inclusión de los denominados *bad outputs* para incorporar indicadores no deseables asociados a la producción que deberían ser también tenidos en cuenta para evaluar el desempeño. Para una revisión de los mismos aplicada al contexto de comparación de sistemas educativos puede acudirse a GIMÉNEZ *et al.* (2017).

(6) Para un estudio acerca de las principales metodologías disponibles para introducir variables de entorno en los análisis de eficiencia puede acudirse a DARAIO y SIMAR (2007); JOHNSON y KOUSMANEN (2011) y SIMAR y WILSON (2011).

(7) A lo largo del artículo utilizo el término genérico de «industria» para referirme a cualquier servicio público que cumpla las características de esta definición.

(8) En Fórmula 1 esto sería equivalente a comprobar que cada año nuestro coche, dando vueltas al mismo trazado, reduce los tiempos. Aunque, sin duda, este cambio parece ir por el buen sentido, esta evaluación resulta vacía sin comparaciones. El siguiente paso sería poner a competir a dos coches de nuestro mismo equipo que nos sirva para incentivar una sana competencia entre ingenieros y pilotos con ánimo de introducir mejoras en los coches. Finalmente, tendremos que compararnos con coches de otros equipos.

(9) Una limitación de esta metodología es que las medidas de diferencias en producti-

vidad pueden variar al cambiar la referencia utilizada. Sin embargo, APARICIO, ORTIZ y SANTÍN (2021) muestran cómo la correlación entre las medidas obtenidas utilizando dos referencias distintas varía entre 0,94 y 0,99. Como ya se ha señalado, idealmente, evaluadores, unidades evaluadas y los usuarios de la evaluación deberían decidir qué tecnología usar como referencia que incluso podría ser la metafrontera en un período base.

(10) Por ejemplo, FELICIANO, LÓPEZ-TORRES y SANTÍN (2021) evalúan el programa Escuela 2.0 comparando la evolución de las diferencias en productividad de los resultados en Educación Secundaria de las CC. AA. que aplicaron el programa respecto a la Comunidad de Madrid que no lo aplicó. Los principales resultados muestran que las CC. AA. que aplicaron el programa perdieron en promedio un 13,3 por 100 de productividad respecto a Madrid.

(11) En cualquier caso, dentro de esta primera ganancia de eficiencia es fácil cuantificar la eliminación de costes duplicados que automáticamente pueden ser reducidos. En el ejemplo hipotético de la fusión de dos departamentos de universidad pequeños, digamos de 20 profesores que pasan a ser un único departamento de 40 profesores, se pasaría de dos directores, dos secretarios académicos y dos personas de administración y servicios a solo una, pudiendo dedicar los recursos liberados a otras tareas.

(12) En mi opinión, una única Agencia Nacional como la que plantean DE LA FUENTE *et al.* (2021) que lleve la evaluación «de arriba hacia abajo» va a encontrar mayores problemas de coordinación y retenciones a la hora de obtener toda la información que si se opta por agencias autonómicas que realicen la evaluación «de abajo a arriba». Algunas de estas agencias podrían nacer por acuerdo para evaluar a varias comunidades autónomas simultáneamente. Con la escasa cultura evaluadora existente en España, una Agencia Nacional será percibida como un organismo fiscalizador y habrá retenciones y retrasos a la hora de compartir con ella la información. Sin embargo, una agencia propia puede contribuir a obtener mejor información y a incidir en que la evaluación solo tiene como objetivo proporcionar información veraz para mejorar los servicios.

BIBLIOGRAFÍA

AIREF (2017). *I Plan de Acción de la revisión del gasto en subvenciones del conjunto de las Administraciones Públicas. (Spending Review)*.

AIREF (2018). *Informe de evaluación de la Sociedad Estatal Correos y Telégrafos, y la prestación del servicio postal universal*.

<p>ÁLVAREZ, A. (Coord.) (2001). <i>La medición de la eficiencia y la productividad</i>. Madrid: Ediciones Pirámide.</p> <p>APARICIO, J., ORTIZ, L. y SANTÍN, D. (2021). Comparing group performance over time through the Luenberger productivity indicator: An application to school ownership in European countries. <i>European Journal of Operational Research</i>, 294(2), pp. 651-672.</p> <p>APARICIO, J. y SANTÍN, D. (2018). A note on measuring group performance over time with pseudo-panels. <i>European Journal of Operational Research</i>, 267(1), pp. 227-235.</p> <p>BATTESE, G. E. y RAO, D. P. (2002). Technology gap, efficiency, and a stochastic metafrontier function. <i>International Journal of Business and Economics</i>, 1(2), pp. 87-93.</p> <p>BERG, S. A., FØRSUND, F. R. y JANSEN, E. S. (1992). Malmquist indices of productivity growth during the deregulation of Norwegian banking, 1980-89. <i>The Scandinavian Journal of Economics</i>, S211-S228.</p> <p>BOGETOFT, P. y OTTO, L. (2010). <i>Benchmarking with DEA, SFA, and R</i> (Vol. 157). Springer Science & Business Media.</p> <p>BOGETOFT, P. y WANG, D. (2005). Estimating the potential gains from mergers. <i>Journal of Productivity Analysis</i>, 23(2), pp. 145-171.</p> <p>CAMANHO, A. S. y DYSON, R. G. (2006). Data envelopment analysis and Malmquist indices for measuring group performance. <i>Journal of Productivity Analysis</i>, 26(1), pp. 35-49.</p> <p>CAVES, D., CHRISTENSEN, L. y DIEWERT, W. (1982). The economic theory of index numbers and the measurement of input, output and productivity. <i>Econometrica</i>, 50(6), pp. 1393-1414.</p> <p>COELLI, T. J., RAO, D. S. P., O'DONNELL, C. J. y BATTESE, G. E. (2005). <i>An introduction to efficiency and productivity analysis</i>. Springer Science & Business Media.</p> <p>COLEGIO DE ECONOMISTAS DE MADRID (2005). La eficiencia de los servicios públicos:</p>	<p>viejos problemas, nuevos enfoques. <i>Revista Economistas</i>, 105. Madrid: Colegio de Economistas de Madrid.</p> <p>CORDERO, J. M., PEDRAJA-CHAPARRO, F. M. y SUÁREZ PANDIELLO, J. (2022). <i>Estudio sobre organización territorial municipal en la provincia de Badajoz</i>. Mimeo. Recuperado de: https://www.donbenito.es/wp-content/uploads/2022/01/INFORME-FINAL-CON-RESUMEN-EJECUTIVO-ESTUDIO-DE-ORGANIZACION-TERRITORIAL-PEDRAJA-PANDIELLO.pdf</p> <p>CORDERO, J. M., SANTÍN, D. y SIMANCAS, R. (2017). Assessing European primary school performance through a conditional nonparametric model. <i>Journal of Operational Research Society</i>, 68(4), pp. 364-376.</p> <p>DARAO, C. y SIMAR, L. (2007). Conditional nonparametric frontier models for convex and nonconvex technologies: a unifying approach. <i>Journal of Productivity Analysis</i>, 28(1), pp.13-32.</p> <p>DE LA FUENTE, Á. (Coord.) (2021). La evaluación de políticas públicas en España: antecedentes, situación actual y propuestas para una reforma. <i>Fedea Policy Papers</i>, 2021-09. Fedea.</p> <p>ESPASA, M. y ESTELLER-MORE, A. (2015). Analyzing judicial courts» performance: inefficiency vs. congestion. <i>Revista de Economía Aplicada</i>, 23(69), pp. 61-82.</p> <p>FARRELL, M. J. (1957). The measurement of productive efficiency. <i>Journal of the Royal Statistical Society: Series A (General)</i>, 120(3), pp. 253-281.</p> <p>FELICIANO, D., LÓPEZ-TORRES, L. y SANTÍN, D. (2021). One Laptop per Child? Using Production Frontiers for Evaluating the Escuela 2.0 Program in Spain. <i>Mathematics</i>, 9(20), 2600.</p> <p>FERREIRA, D. y MARQUÉS, R. C. (2015). Did the corporatization of Portuguese hospitals significantly change their productivity? <i>The European Journal of Health Economics</i>, 16(3), pp. 289-303.</p> <p>FERREIRA, D. C., MARQUES, R. C. y PEDRO, M. I. (2016). Comparing efficiency of holding business model and individual management model of</p>	<p>airports. <i>Journal of Air Transport Management</i>, 57, pp.168-183.</p> <p>FRIED, H. O., LOVELL, C. K. y SCHMIDT, S. S. (Eds.). (2008). <i>The measurement of productive efficiency and productivity growth</i>. Oxford University Press.</p> <p>FUNCAS (2003). Sector público y eficiencia. <i>Papeles de Economía Española</i>, 95. Madrid: Funcas.</p> <p>GIMÉNEZ, V., THIEME, C., PRIOR, D. y TORTOSA-AUSINA, E. (2017). An international comparison of educational systems: a temporal analysis in presence of bad outputs. <i>Journal of Productivity Analysis</i>, 47(1), pp. 83-101.</p> <p>GIMÉNEZ, V., THIEME, C., PRIOR, D. y TORTOSA-AUSINA, E. (2019). Comparing the performance of national educational systems: inequality versus achievement? <i>Social Indicators Research</i>, 141(2), pp. 581-609.</p> <p>JOHNSON, A. L. y KUOSMANEN, T. (2011). One-stage estimation of the effects of operational conditions and practices on productive performance: asymptotically normal and efficient, root-n consistent StoNEZD method. <i>Journal of Productivity Analysis</i>, 36(2), pp. 219-230.</p> <p>KUMBHAKAR, S. C., WANG, H. y HORNCastle, A. P. (2015). <i>A practitioner's guide to stochastic frontier analysis using Stata</i>. Cambridge University Press.</p> <p>LEIBENSTEIN, H. (1966). Allocative efficiency vs. «X-efficiency». <i>The American Economic Review</i>, 56(3), pp. 392-415.</p> <p>LEVIN, H. M. (1974). Measuring efficiency in educational production. <i>Public Finance Quarterly</i>, 2(1), pp. 3-24.</p> <p>LEVIN, H. M., JAMISON, D. T. y RADNER, R. (1976). Concepts of economic efficiency and educational production. <i>Education as an Industry</i>, pp. 149-198. NBER.</p> <p>MAS-COLELL, A., WHINSTON, M. D. y GREEN, J. R. (1995). <i>Microeconomic theory</i> (Vol. 1). New York: Oxford University Press.</p> <p>MERGONI, A. y DE WITTE, K. (2022). Policy evaluation and efficiency: a systematic literature</p>
--	--	---

<p>review. <i>International Transactions in Operational Research</i>, 29(3), pp. 1337-1359.</p> <p>MEFP (2021). <i>Anuario estadístico: Las cifras de la educación en España. Edición 2021 (Curso 2018-2019)</i>. Madrid: Ministerio de Educación y Formación Profesional.</p> <p>MINISTERIO DE SANIDAD (2021). <i>Informe Anual del Sistema Nacional de Salud 2019. Aspectos destacados</i>. Madrid: Ministerio de Sanidad.</p> <p>NARBÓN-PERPIÑÁ, I. y DE WITTE, K. (2018a). Local governments» efficiency: a systematic literature review – part I. <i>International Transactions in Operational Research</i>, 25(2), pp. 431-468.</p> <p>NARBÓN-PERPIÑÁ, I. y DE WITTE, K. (2018b). Local governments» efficiency: a systematic literature review – part II. <i>International Transactions in Operational Research</i>, 25(4), pp. 1107-1136.</p>	<p>NISKANEN, W. A. (1975). Bureaucrats and politicians. <i>The Journal of Law and Economics</i>, 18(3), pp. 617-643.</p> <p>OATES, W. E. (1999). An essay on fiscal federalism. <i>Journal of Economic Literature</i>, 37(3), pp. 1120-1149.</p> <p>OH, D. H. (2010). A global Malmquist-Luenberger productivity index. <i>Journal of Productivity Analysis</i>, 34(3), pp. 183-197.</p> <p>PASTOR, J. T. y LOVELL, C. K. (2005). A global Malmquist productivity index. <i>Economics Letters</i>, 88(2), pp. 266-271.</p> <p>PEDRAJA-CHAPARRO, F. M., SALINAS-JIMÉNEZ, J. y SMITH, P. (1999). On the quality of the data envelopment analysis model. <i>Journal of the Operational Research Society</i>, 50(6), pp. 636-644.</p> <p>SANTÍN, D. Y SICILIA, G. (2017). Impact evaluation and frontier methods in education: a step forward. En:</p>	<p>G. JOHNES, J. JOHNES, T. AGASISTI y L. LÓPEZ-TORRES (eds.), <i>Handbook of Contemporary Education Economics</i>. Cheltenham: Edward Elgar Publishing.</p> <p>SCRGSP, <i>Steering Committee for the Review of Government Service Provision (2013). Report on Government Services 2013</i>. Canberra: Productivity Commission.</p> <p>SICKLES, R. C. y ZELENYUK, V. (2019). <i>Measurement of productivity and efficiency</i>. Cambridge University Press.</p> <p>SIMAR, L. y WILSON, P. W. (2011). Two-stage DEA: caveat emptor. <i>Journal of Productivity Analysis</i>, 36(2), pp. 205-218.</p> <p>TIROLE, J. (1994). The internal organization of government. <i>Oxford Economic Papers</i>, 46, pp. 1-29.</p>
---	--	---