REVISTA VASCA DE ECONOMÍA 60

EKONOMIAZ

III · 2005

La evaluación de las políticas públicas

EKONOMIAZ

N.º 60 - 3er CUATRIMESTRE - 2005

VOLUMEN I

LA EVALUACIÓN DE LAS POLÍTICAS PÚBLICAS

EUSKO JAURLARITZA GOBIERNO VASCO

OGASUN ETA HERRI ADMINISTRAZIO SAILA

DEPARTAMENTO DE HACIENDA Y ADMINISTRACIAÓN PÚBLICA

Eusko Jaurlaritzaren Argitalpen Zerbitzu Nagusia

Servicio Central de Publicaciones del Gobierno Vasco

Vitoria-Gasteiz, 2006

CONSEJO DE REDACCIÓN / ERREDAKZIO KONTSEILUA

Iñaki Beristain Etxabe (Director)
Javier Gúrpide Ibarrola (Subdirector)
Alberto Alberdi Larizgoitia, Carlos Agustín Alvarado García, Jesús Ferreiro Aparicio, Juan Antonio Maroto Acín, Mikel Navarro Arancegui, Juan Fco. Santacoloma Sanz, Fco. Javier Sasigain Saenz de Buruaga,

REDACCIÓN Y ADMINISTRACIÓN / ADMINISTRAZIO ETA ERREDAKZIO

Departamento de Hacienda y Administración Pública/Ogasun eta Herri Administrazio Saila
c/ Donostia-San Sebastián, s/n.
01010 Vitoria-Gasteiz
Teléfonos: 945/01 90 38. Administración
945/01 90 36. Redacción

http://www1.euskadi.net/ekonomiaz

SUSCRIPCIÓN ANUAL / URTEKO HARPIDETZA:

Particulares: 18 €. Instituciones y empresas: 30 €.
Precio de un ejemplar/Ale bakarren prezioa: 12 €.

EKONOMIAZ aldizkarian parte hartu nahi dutenek Eusko Jaurlaritzako Ogasun eta Herri Administrazio Sailera igorri beharko dituzte beren idazlanak.

Erredakzio Kontseilua beretzat gordetzen du jasotako artikuluak argitaratzeko eskubidea, aurrez ebaluatzaile anonimo baten edo batzuen iritzia ezagutuko duelarik. Idazlana argitaratzeak ez du, berez, edukiarekiko adostasunik adierazi nahi, beraren erantzunkizuna egilearena eta ez beste inorena izango delarik.

Artikuluak aurkezteko arauak honako web orri honetan daude: http://www1.euskadi.net/ekonomiaz/articulos_e.asp

EKONOMIAZ aldizkariko artikuluak indexatuak dira ondoko datutegietan: CSICeko ISOC eta American Economic Association elkarteko *Journal of Economic Literature*-k argitaratzen dituen JEL CD formatoan, e-JEL eta EconLit.

Todas aquellas personas que deseen colaborar en EKONOMIAZ deberán enviar sus artículos al Departamento de Hacienda y Administración Pública del Gobierno Vasco.

El Consejo de Redacción se reserva el derecho de publicar los artículos que reciba, previo sometimiento a un proceso de evaluación anónima doble. Su publicación no significa necesariamente el acuerdo con el contenido que será de responsabilidad exclusiva del autor.

Las normas de presentación de los artículos se pueden encontrar en la página web: http://www1.euskadi.net/ekonomiaz/articulos_c.asp

Los artículos de EKONOMIAZ son indexados en las bases de datos ISOC del CSIC y JEL en CD, e-JEL y EconLit del *Journal of Economic Literature* de la American Economic Association.

Argitaratzailea:	Eusko Jaurlaritzaren Argitalpen Zerbitzu Nagusia
Edita:	Servicio Central de Publicaciones del Gobierno Vasco Donostia-San Sebastián, 1 - 01010 Vitoria-Gasteiz
Fotokonposaketa: Fotocomposición:	Gráficas Varona, S.A. Polígono “El Montalvo”, parcela 49. 37008 Salamanca
Imprimaketa: Impresión:	Gráficas Varona, S.A. Polígono “El Montalvo”, parcela 49. 37008 Salamanca
Lege Gordailua: Depósito Legal:	VI. 285-2006

ISSN-0213-3865

SUMARIO

PRESENTACIÓN

LA EVALUACIÓN DE LAS POLÍTICAS PÚBLICAS

Coordinador: **Guillem López Casasnovas**

Catorce puntos esenciales sobre la evaluación de las políticas públicas con especial referencia al caso de las políticas sociales	**18**	*Joan Subirats Humet*
Evaluación y gestión eficiente en el sector público: aspectos organizativos e institucionales	**38**	*Jorge Onrubia Fernández*
De la evaluación a la gestión: acortar la brecha entre eficacia y efectividad	**58**	*Vicente Ortún Rubio*
Programación y evaluación pública: un triángulo complejo	**76**	*José Luis Osuna Llaneza* *Carmen Vélez Méndez* *Ana Cirera León* *Juan Murciano Rosado*
Principales mecanismos de evaluación económica de políticas públicas	**98**	*Ivan Planas Miret*
Métodos cuantitativos y "Benchmarking": su utilidad para orientar las políticas públicas	**122**	*Beatriz González López-Valcarcel*
Algunas reflexiones metodológicas sobre la evaluación de la eficiencia productiva de las instituciones de educación superior	**140**	*José María Gómez Sancho* *María Jesús Mancebón Torrubia*
El uso del análisis coste-beneficio en la estimación de la posibilidad de rescate de las autopistas	**168**	*Albert Castellanos Maduell* *Guillem López Casasnovas*
Estimación del valor monetario de los años de vida ajustados por calidad: estimaciones preliminares	**192**	*José Luis Pinto Prades* *Jorge-Eduardo Martínez Pérez*
Política europea de empleo para mujeres inactivas: ¿cómo sabremos si estamos acertando?	**210**	*Xavier Ballart Hernández*
El margen de maniobra presupuestario de las Comunidades Autónomas. Método de análisis y aplicación a la Comunidad Autónoma de Aragón	**232**	*Ramón Barberán Ortí*

Sumario

ZORIONAK!

Veinte años no son nada, y menos en la dilatada y compleja historia de un pueblo milenario como el nuestro. Y, sin embargo, ¡cómo ha cambiado Euskadi en los últimos veinte años! Nada mejor que una publicación como *Ekonomiaz* para atestiguarlo.

Una revista que nació con una doble vocación que hasta la fecha ha logrado cumplimentar con creces.

Ekonomiaz ha sido y es, por una parte, un vehículo para la generación, transmisión y divulgación de conocimientos económicos, vehículo que siempre ha estado a disposición de la comunidad científica precisamente para que todos y todas fuéramos partícipes de tales conocimientos.

En segundo término, y de manera complementaria a la primera, la revista ha mantenido siempre presente su vocación de referente en la adopción de las decisiones de las Administraciones públicas vascas, fundamentalmente en el área de la política económica —entendida ésta en su sentido más amplio—. No debe perderse de vista que el surgimiento de *Ekonomiaz* tuvo lugar cuando la profunda crisis que azotaba la economía vasca obligó a nuestras instituciones de autogobierno a llevar a cabo apuestas activas y decididas en el proceso de reestructuración y reconversión del tejido productivo. Desde esta perspectiva, la revista desempeñó —y lo sigue haciendo— un papel clave a la hora de proporcionar esas herramientas de análisis y reflexión que coadyuvan al diseño, la adopción y hasta la evaluación de las políticas económicas más pertinentes y eficaces.

Sumario

Junto a esta doble vocación, la publicación ha mantenido, además, una visión penetrante, mediante la cual ha sabido conjugar el análisis de los asuntos de más candente actualidad, las reflexiones de tipo más estructural o a largo plazo, y las experiencias ceñidas al terreno académico de nuestros economistas y siempre con las aportaciones de los estudiosos de fuera de Euskadi. Actualidad, visión a largo plazo y universalidad sin perder de vista el propio enraizamiento constituyen, desde esta perspectiva, las señas de identidad de ***Ekonomiaz***, un cóctel que, a mi juicio, explica a la perfección el éxito que a lo largo del tiempo viene cosechando.

En efecto, la elección de las temáticas que han dado contenido a la revista siempre ha estado guiada por un criterio de oportunidad en el sentido de optar por aquellas cuestiones que en cada momento se hallaban en el centro del debate académico, social y político.

Ahora bien, ello no ha supuesto ningún límite a la hora de concebirla como una herramienta de prospección a medio y largo plazo, tratando de situarse lo más cerca posible de la vanguardia del conocimiento económico planteando temas que se intuía iban a ser objeto de especial interés en el futuro. Baste señalar, a modo de botón de muestra de lo que digo, que en los cinco primeros años de su edición —allá por la segunda mitad de la década de los ochenta— ***Ekonomiaz*** ya había analizado materias que hoy en día están configurando el pensamiento socioeconómico más avanzado, como pueden ser el papel de la educación y la formación —del capital humano— en los procesos de crecimiento y desarrollo económicos, la importancia capital de la investigación y la innovación tecnológica en tanto que factores de competitividad, o, por no extenderme más, la salvaguarda del medio ambiente y la racionalidad en el uso de los recursos naturales como elementos fundamentales de un desarrollo sostenible.

La modernidad de estos asuntos resulta más que evidente, de igual manera que es palpable el acierto de la revista al anticipar la discusión en torno a los mismos con veinte años de antelación.

He resaltado, asimismo, que al recabar la colaboración de los especialistas, ***Ekonomiaz*** no se ha circunscrito, ni muchísimo menos, a su ámbito territorial natural, sino que, al contrario, siempre ha tratado de abrirse a una dimensión lo más amplia posible. No estamos ante una publicación local que sólo se ha hecho eco de las opiniones, análisis y reflexiones de los economistas de Euskadi en relación a cuestiones de interés fundamental o exclusivamente vascos. Así, con el paso de los años en sus páginas ha ido creciendo paulatinamente la participación de expertos extranjeros, cuyas aportaciones han definido, en numerosas ocasiones, la frontera misma de la ciencia económica del momento.

La combinación de todos los ingredientes citados ha arrojado como resultado final un producto de calidad, de reconocido prestigio no sólo en Euskadi sino más allá de nues-

Sumario

tro propio ámbito, en definitiva, una pieza esencial en la labor desarrollada por el conjunto de Administraciones públicas vascas.

Por ello me he querido sumar, por medio de estas líneas, al merecido homenaje que debemos tributar a todos y todas quienes han hecho y hacen posible esta realidad que cumple ya veinte años. Zorionak!

Idoia Zenarrutzabeitia Beldarrain

VICELEHENDAKARI
CONSEJERA DE HACIENDA Y ADMINISTRACIÓN PÚBLICA
GOBIERNO VASCO

Sumario

Sumario

PRESENTACIÓN

LAS BRECHAS QUE VAN DEL ANÁLISIS A LA DEFINICIÓN DE LA POLÍTICA PÚBLICA; DE LA IMPLEMENTACIÓN DE LA DECISIÓN A SU EVALUACIÓN; DE LA EFICACIA TEÓRICA A LA EFECTIVIDAD CONCRETA EN LA GESTIÓN PÚBLICA

DEL «QUÉ» Y DEL «CÓMO» DE LAS POLÍTICAS PÚBLICAS Y DE SU EVALUACIÓN

Muchas veces, aunque no esté suficientemente asumido por buena parte de los estudiosos de la economía, es la mano visible del poder público y no la mano invisible *smithiana* la que maneja, interfiere, reasigna, en su conjunto, casi la mitad de toda la renta y riqueza generada por el país en un momento dado. El mercado crea riqueza pero el sector público la reasigna en buena parte entre los destinatarios. Esto es al menos lo que refleja el creciente peso del gasto consolidado de las administraciones públicas en España, al que cabría añadir, por lo demás, el resultante del resto de la intervención pública no administrativa en la economía, por la vía de la regulación impuesta al sector privado, y del sector público empresarial. Todo ello se traduce, más allá de las coyunturas de déficit o de superávit público en la generación de ingresos extraordinarios —por la vía por ejemplo de las privatizaciones—, una creciente presión fiscal soportada por el conjunto de la economía.

Las “cuñas” fiscales con ello creadas, al afectar a los costes, inciden en la competitividad internacional de nuestras mercancías y servicios para la exportación, y por la vía de la inflación general en los salarios y remuneraciones reales de la población. Algunos aducirán que los niveles de presión fiscal y gasto público españoles, en comparación

con los de otros países occidentales, no son relativamente altos. Olvidan éstos, sin embargo, que en economías de alto rendimiento es la dinámica de precios la que influye en la competitividad y no los niveles relativos. Y así, en los últimos veinte años de la economía española, el crecimiento de presión fiscal bate records entre los países de la OCDE. En cualquier caso, para el lector escéptico, sea alta o baja la presión fiscal, no debe quedarle duda alguna, de que un sector público eficiente es una plataforma para el crecimiento económico, y un sector público ineficiente, su mayor lastre.

De todo lo anterior se deduce mi interés por la eficiencia de las actuaciones públicas en la economía. Debo ser en este sentido uno de los pocos economistas que partiendo de la más ortodoxa formación académica parece hoy tener «vendida su alma al diablo» de las políticas y politólogos. Dicha «traición» es mayor aún en mi caso, ya que el tránsito recorrido fluye de la economía pública hasta los umbrales de las políticas públicas, mucho más ligadas éstas a la Política a secas. Ello contrasta con el papel de otros macro-micro economistas que con preocupación y tareas similares fundamentan desde aquellas bases de la economía convencional, la denominada Economía Política (Political Economy, no confundir aquí con Política económica, área tradicionalmente moderada de la misma Economía Aplicada), mucho más aséptica y menos políticamente «contaminada».

Continuando el relato personal, como auto justificación de quien suscribe y justificación del monográfico que el lector tiene ahora en sus manos, y en respuesta a la amable invitación que en su día me hizo el Consejo Editorial de Ekonomiaz, tengo el convencimiento de que la trasgresión de la Economía a la Política pública no es un fruto tardío de los muchos años de ejercicio profesional o del debilitamiento neuronal, sino que valía la pena en sí misma. Dicho convencimiento proviene desde hace ya mucho tiempo, de quien como a mí le toca cada año impartir entre sus alumnos las lecciones de la Teoría de la Hacienda Pública relativas al «fallo» del mercado y a las ventajas «pareto eficientes» de la intervención pública correctora. O cuando, siguiendo el segundo teorema de la economía del bienestar, trata de demostrar cómo el sector público puede variar la dotación inicial de recursos o corregir asignaciones que se adecuen en mayor medida a las preferencias sociales. Estoy convencido de que tan importante resulta centrar la enseñanza en justificar lo primero (la intervención pública) como en garantizar lo apropiado de esta misma intervención al servicio de los objetivos definidos socialmente. En este campo cualquier observador puede detectar una enorme brecha entre lo que es la literatura económica convencional del estudio de las razones, naturaleza, propósito, etc. de la intervención pública; esto es, acerca de cuál ha de ser el papel del sector público en la economía, y el poco interés que suscita entre economistas el «cómo» de dicha intervención. Como si de un tema de derecho administrativo se tratase, se olvida que sin ello, no se puede garantizar que la ineficiencia producida por la intervención pública no acabe siendo mayor que la pérdida de eficiencia de los “fallos” de mercado que persigue recuperar. Y ello pese a aceptar, como sin duda se hace en nuestra profesión, la importancia de la eficiencia pública en la competitividad internacional de una economía (cuñas fiscales, condicionantes infraestructurales, presión fiscal relativa...), tal como hemos comentado más arriba, en particular en economías de alto rendimiento.

Sumario

VA "DE BRECHAS"

En este monográfico vamos a hablar «de brechas». Entre la economía normativa y la positiva; entre disciplinas que se impermeabilizan entre sí para así garantizar su «cientifismo» (incluso a través de su "nomenclátor endogámico protector", con escasas sociedades científicas interdisciplinares y ausencia de publicaciones sensibles a la materia —de ahí mi interés por la iniciativa de *Ekonomiaz*), como acabo de comentar. Brechas entre el consenso teórico y la práctica político administrativa en la propia evaluación de las políticas públicas. Como señala en su texto Joan Subirats, el *gap* entre lo que se acepta como bueno —y se debería por tanto hacer en materia de evaluación de políticas— y la realidad más acomodaticia es enorme. Digan lo que digan los textos legislativos, sea cual sea la restricción impuesta sobre el gasto social, proclamen lo que proclamen los políticos, la inercia administrativa siempre acaba dominando. Existe brecha por lo demás entre la eficacia de lo que se sabe que funciona (gestión basada en la evidencia) y la efectividad de su seguimiento en el caso concreto (incentivos a aplicar lo que se sabe y que los gestores resuelvan los problemas posteriores que se presentan) tal como muestra con clarividencia Vicente Ortún en la gestión sanitaria más clínica. En resumen, brechas entre la Economía y la Política, entre la Economía y la Política pública, entre el diseño de las intervenciones y su evaluación, y entre éstas y la propia gestión pública.

CAMBIO ORGANIZATIVO, CONDICIÓN NECESARIA NO SUFICIENTE

El resto del Monográfico acompaña el objetivo antes comentado. Los encargos efectuados pretenden, por un lado, enmarcar los condicionantes para una evaluación y gestión eficientes en el sector público en distintos aspectos organizativos e institucionales (Jorge Onrubia); por el otro, identificar aquellos procesos facilitadores para una utilización eficiente de la información como base del funcionamiento adecuado de los sistemas de control interno y externo de las políticas (José Luis Osuna). Así, se trata de comparar dos modelos polares de arquitectura organizativa, uno descentralizado, con autonomía en la gestión operativa, y otro centralizado y con jerarquía en la adopción de decisiones, y valorar *pros* y *contras* de ambos modelos en cuanto a requerimientos de mecanismos e incentivos para el desempeño y consecución de una gestión pública más eficiente. Parece evidente que sin atender al cambio organizativo, las propuestas reformadoras de la actuación pública están llamadas a fracasar, ya que —entre otras razones— no se valoran como «apropiables» (referenciables al cambio inducido) por los propios agentes que las han de llevar a cabo. Para ello una arquitectura descentralizada y con plena autonomía operativa de sus unidades gestoras (agencias o estructuras «divisionales» frente a las «departamentales») ha de resultar inequívocamente superior debido a las ventajas que supone desde el punto de vista del uso de la información (asociables entre otras a los contratos de base presupuestaria entre agencias y financiador público). Es relevante, finalmente, una mayor delimitación entre responsabilidades de índole política y técnica, ya que la permeabilidad elevada entre políticos y funcionarios acaba suponiendo, como muestra el caso español, un freno a las reformas organizativas más eficaces y a aceptar la evaluación (aunque sea por una simple comparación).

Sumario

PUENTES TRANSITABLES E INTRANSITABLES

Los puentes a transitar entre la Economía —la teoría de la organización en este caso— y la Política no están exentos de controversia. Qué hay de común en el paradigma Administración y Empresa, cuáles son las zonas comunes y los rasgos distintivos, es objeto de debate en lo que haya de imitable y lo que sea simplemente emulación equívoca. Trabajar a partir de la cooperación y la interdependencia o competir. Especializarse en nichos de ventaja competitiva o poner a disposición común el conjunto de servicios para solucionar «casos». Qué hay de artesanal y en qué ámbitos cabe aplicar la lógica industrial, cuando la efectividad (y eficiencia con la incorporación de costes) de lo que se hace se orienta eficazmente a lo que se había definido como objetivo último (no intermedio). Se puede ser más eficiente especializándose, pero no se es eficaz si la geografía exige mantener una oferta completa. Se reducen sin duda los costes unitarios aumentando la actividad, pero aumentar la actividad puede no ser el objetivo (más bien prevenir y hacer innecesarias aquellas otras actuaciones puede ser la finalidad). Servicios valorados por su disponibilidad, *stand by* de oferta, como externalidad de opción o cobertura «en caso de», y no para mercantilizar su utilización (a través de técnicas de marketing). La confusión entre objetivos y medios: seleccionar adversamente los casos para mostrar eficiencia en costes frente a atender y afrontar los casos de riesgo más alto (en sanidad, en educación compensatoria) pese a su mayor impacto presupuestario. Y podríamos seguir con otros muchos ejemplos.

LA SALVAGUARDA DE LO COMPLEJO

Sin embargo, reconocido ciertamente lo anterior, no vale ampararse en la ambigüedad de los objetivos políticos, en la supuesta imposible separación entre política pública y gestión, entre el perfil político y el del trabajador por «cuenta ajena» de la administración, o escudarse en la «completa heterogeneidad» de los supuestos contemplados como para no permitir comparación alguna (*benchmarking*). Es cierto que en las políticas públicas confluyen distintos agentes con intereses contrapuestos, con un complejo proceso de formación de decisiones. Pero no todos los intereses son igualmente legítimos. La ambigüedad escuda conductas y favorece consensos, pero a unos les corresponde definir objetivos (es su obligación, o al menos su tarea) y a otros ejecutarlos. Para unos el gobierno y la gobernanza de las situaciones, para otros la gestión. Para unos es más el énfasis en lo estratégico que en lo operacional, y viceversa para otros. La evaluación es diferente, en consecuencia, en ambos casos. La primera supone una criba más social y política (electoral finalmente); la segunda tiene más componentes técnicos, logísticos y operativos.

La empresa privada no es la Administración pública, pero la adecuación de fines y medios es consubstancial a toda organización, la Administración inclusive. Precisamente porque la legitimación social está sobre todo en la política más que en la gestión conviene separar los dos ámbitos, para que no se beneficie el gestor por la vía política, ilegíti-

mamente de su ineficiencia. Sin duda la realidad requiere matices; pero conviene ahora hacer hincapié en ello para evitar un cierto «totum revolutum» existente en la nueva gestión pública. Al gestor se le ha de exigir formación y «expertise», flexibilidad en la necesaria cooperación interadministrativa, en el trabajo en red, etc., y adecuación al cambio político (contra el gestor que disfraza su comodidad de mesianismo político a coste cero) y al interés del usuario, dentro de las restricciones impuestas por la Administración que está política y socialmente legitimada. Contravengo pues parcialmente los argumentos de uno de los autores del Monográfico cuando afirma que «el debate sobre el papel y los resultados de la acción pública no se sitúa ya sólo en el cómo hacer las cosas, sino que exige contar con instrumentos que permitan responder asimismo a las preguntas de a qué se deben ocupar los poderes públicos y de quién puede ofrecer mejores resultados eficaces y eficientes...». Para este observador, la cuestión del cómo hacer las cosas continúa siendo fundamental...Por higiene gestora, antes de pasar a la evaluación de segunda y tercera generación, conviene garantizar la más básica: la de qué hacemos, cómo lo hacemos y para qué sirve. Ello es especialmente importante cuando los objetivos se establecen de modo coherente con una descentralización operativa efectiva y se insertan incentivos a la coordinación, por ejemplo en materias como la exclusión social o la lucha contra la pobreza. En el análisis de sus resultados, la ambigüedad política es aquí fácilmente resoluble: se disminuyen o no los afectados, se consigue o no con efecto permanente..., no cabiendo aquí demasiado debate social y político.

¿SON COMPARABLES LOS TRAJES A MEDIDA?

No hay duda que la multidimensionalidad de las acciones para objetivos como los formulados (contra la exclusión, la inserción laboral...) es muy importante, como son relevantes tanto las trayectorias personales como los condicionantes objetivos. Ello exige por tanto creatividad e innovación en la gestión pública, como factores clave en la definición de políticas heterogéneas, flexibles y tendencialmente personalizadas. Los mecanismos decisionales han de favorecer dicha heterogeneidad y adaptación al caso concreto, y el trabajo posterior en red, pero también su evaluación, finalmente, sobre el análisis del éxito o no en su consecución, permitiendo el aprendizaje al menos de los factores condicionantes de los resultados en uno u otro sentido.

Parece por tanto crucial la forma de insertar la evaluación en los procesos decisionales por cuanto debe proporcionar información objetiva a los gestores de los programas y a los representantes políticos sobre los costes y los efectos de los programas. Destaca en este sentido el estudio, en varios trabajos de esta monografía, de los aciertos y errores de la política europea de evaluación —políticas intergubernamentales de nueva estirpe— de los programas agrícolas, de desarrollo regional, empleo y formación entre otros, constitutivos de nuevas áreas de desarrollo de políticas públicas impulsadas por la Unión Europea. Todo apunta a que existen claros riesgos de excesiva institucionalización de los procedimientos de evaluación hoy seguidos (por ejemplo, en la experiencia evaluadora de los Fondos estructurales en el País Vasco), entendidos como justificación *ex post*

de decisiones tomadas y que, lejos de aportar valor, acaban siendo incluso obstáculo para su finalidad indirecta de instrumento de aprendizaje y revisión de la determinación de las propias políticas. En concreto, para el caso de los programas de ocupación de mujeres inactivas Xabier Ballart señala que la experiencia pone de relieve que la evaluación estandarizada seguida, no responde a preguntas fundamentales sobre el valor público del programa. Parece pues que la nueva evaluación europea incide en los mismos errores que la tradicional, al ser incapaz de adaptarse a las realidades de cada proyecto. Ciertamente sirve para comparar indicadores básicos entre programas (actividades, recursos empleados, algunos resultados parciales), pero difícilmente tiene papel alguno para fundamentar mejor las decisiones que hay que tomar (más allá de garantizar su corrección) y sobre todo, para revisar las estrategias de las distintas políticas emprendidas (más personalizadas, flexibles y temporales). Para ello se requiere a menudo una información más cualitativa que la que se demanda, menos rutinaria, anodina y burocratizada, capaz de propiciar debates de mayor relevancia teórica y política.

LOS INCENTIVOS A LA TOMA EN CONSIDERACIÓN

Cualquier tipo de evaluación proporciona conocimiento analítico, y aunque siempre se pueda conocer, investigar y evaluar algo más, el auténtico déficit radica en la puesta en práctica de lo que en general ya se conoce y que la evaluación a menudo ratifica. El quid está en el talento implantador para convertir todo ello en mejor gestión. A veces el gestor carece del conjunto de incentivos para pasar de la teoría a la práctica. En este caso identificar las restricciones de entorno como exógenas y políticas y tratarlas de modo diferente a las endógenas mitigables a través de la acción, etc. resulta prioritario; por ejemplo, favoreciendo una mejor organización, coordinación, etc., tal como señala Vicente Ortún.

En la gestión presupuestaria concreta, la corrección del «dicho» al «hecho» requiere unos mínimos de discrecionalidad en las actuaciones que el presupuesto a menudo, tal como lo tenemos hoy formulado, no permite. Véase Ramón Barberán, por ejemplo, para el cálculo empírico de los márgenes potenciales y netos, y los márgenes de maniobra efectivamente utilizados, a los efectos derivados de variar a corto plazo las políticas de gasto ya en marcha y dar inicio a otras nuevas. Existen además, condicionantes interadministrativos que restringen la libertad de asignación de recursos entre políticas mediante la concesión de transferencias financieras condicionadas. Por tanto, merece ser muy cuidadoso en los consensos políticos creados, ya que a menudo los márgenes de discrecionalidad en la variación de políticas, de la fuerte inercia administrativa, son tan bajos que no existen nuevos márgenes de error.

EL CONOCIMIENTO ES CARO ¿PROBAMOS CON LA IGNORANCIA?

Llevar a cabo cambios en las políticas existentes o introducir nuevas estrategias exige el mejor conocimiento, a ser posible de decisiones basadas en la evidencia disponible.

Sumario

Dicho recurso implica el uso de instrumentos que coadyuven a dichos cambios. En ningún caso se trata de sustituir la decisión política responsable en la materia, sino de informar mejor acerca de sus consecuencias. De ahí que la gestión y evaluación de las políticas públicas no se puede permitir viajes sin alforjas, Ivan Planas destaca en este sentido una exhaustiva panoplia de instrumentos micro y macro, desde los modelos de equilibrio general calibrados a la evaluación coste-beneficio utilizado en diversos ámbitos de la actividad pública, ya sea en materia de grandes infraestructuras Guillén López y Albert Castellanos ya sea en políticas micro muy concretas, como en el caso de los programas de salud buco dental, analizados por Pilar García (en concreto, el plan de atención dental infantil en el País Vasco, por medio del análisis de comparación de dobles y triples diferencias, según las referencias a los grupos de control tomados) recogido en el segundo volumen de este Monográfico. La distinta robustez de los métodos al uso para la construcción de indicadores de eficiencia en la gestión, ya sea para su análisis estático relativo, o dinámico —de mejora, requiere un segundo estudio de detalle (Beatriz González). Se trata cuando menos de adquirir conciencia sobre la conveniencia y alcance —pero también las limitaciones de las distintas técnicas cuantitativas utilizadas, y acerca de la necesidad de acompañamientos cualitativos en su utilización tanto en la evaluación de la gestión de los servicios públicos, como a la hora de matizar su aplicación a cada campo concreto (en sanidad -José Luis Pinto; y en la universidad -José María Gómez y María Jesús Mancebón). Todo ello exige valentía ya que las dificultades son muchas y los márgenes de actuación escasos. De ahí que exista una cierta desidia en su implementación, bien revestida además por tabúes de connotaciones poco democráticas: ¿quién se arriesga a definir el valor monetario de una vida?; ¿a quién le asusta una descomposición de indicadores de inputs, outputs y outcomes?; ¿por qué razón no se levantan barreras a corto para hacer posibles entradas, si así se estima, más a medio y largo plazo?; ¿a quién favorecen los temores del inmovilismo ante el cambio?

En cualquier caso, aún siendo conscientes de las limitaciones de la evaluación económica para los cierres de brecha aquí comentados —desde el análisis a la reelaboración de políticas, de la política pública a su implementación, de la eficacia a la efectividad gestora... para la toma de decisiones públicas—, se nos antoja que las alternativas (la no evaluación, su no utilización) pueden acabar siendo mucho peores. Por lo tanto, si estamos de acuerdo en que elaborar sobre la evidencia para cerrar las brechas resulta caro e incierto, ¿alguien realmente sensato desea hacerlo sobre la base de la ignorancia?

Guillem López Casasnovas
Universitat Pompeu Fabra

Sumario

Catorce puntos esenciales sobre evaluación de políticas públicas con especial referencia al caso de las políticas sociales[1]

En este artículo se reflexiona sobre la utilidad, la calidad y la idoneidad de los instrumentos de evaluación de políticas públicas en el contexto actual. Se explica cómo no pueden limitarse a ser meros controladores de la legalidad o a detectar abusos de poder. Deben ir más allá y convertirse en indicadores reales de impacto además de reflejar la calidad de las interrelaciones entre los organismos públicos y el exterior. Para ello deberían previamente segmentar y establecer parámetros adecuados según el tipo de relación y de estructuras de poder que subyacen a los distintos marcos de actuación de las administraciones públicas. Por otra parte, se analizan las políticas sociales que se han ido poniendo en marcha en España y Europa, y se apuntan cuales deberían ser las claves de una evaluación que sepa dar respuesta a los nuevos escenarios económicos, sociales, familiares y de participación política, caracterizados por la exclusión, la desigualdad y la llamada nueva gobernanza.

Artikulu hau politika publikoak ebaluatzeko tresnek egun duten erabilgarritasuna, kalitate eta egokitasunari buruzko hausnarketa bat da. Bertan azaldu egiten da legaltasunaren kontrolatzaile hutsak izatera edo nagusikeriak antzematera ezin direla mugatu. Haratago jo behar dute eta, erakunde publikoen eta kanpoaldearen arteko harremanen kalitatea islatzeaz gain, eraginaren benetako adierazleak bihurtu behar dira. Horretarako, aldez aurretik, herri administrazioen jarduketa-esparruen azpian dauden harreman eta botere egitura moten araberako parametro egokiak segmentatu eta ezarri beharko lituzkete. Horrez gain, Espainian eta Europan martxan jartzen ari diren gizarte-politikak aztertu egiten dira eta baztertze, desberdintasuna eta delako gobernantza berria ezaugarri dituzten ekonomia, gizarte, familia eta politikan parte-hartzeko egoera berriei erantzuna eman diezaien ebaluaketaren gakoak zeintzuk izan beharko liratekeen adierazten dira.

This article gives some reflections about the usefulness, the quality and the suitability of the instruments of the public policies evaluation in the present context. It explains how they can not limit themselves being just mere controllers of the legality or to detect power abuse. They have to go beyond this, to convert into impact real indicators, and to reflect the quality of the interrelations between the public agencies and outside. For this reason they, previously, have to segment and to establish adequate parameters depending on the type of the relation and on the power structures that underlie to the different fields of the interventions of the public administrations. On the other hand, social policies will be analysed, these ones have been employed in Spain and in Europe. And it will be appointed to the ones that must be the clues to an evaluation that knows to give answers to new economic, social, family scenes and of a political participation characterised by the exclusion, inequality and by the so called new governance.

[1] Este texto reproduce, parcialmente, presentaciones efectuadas por el autor en el Congreso sobre Evaluación de la eficacia en las administraciones públicas, organizado por el Tribunal Vasco de Cuentas Públicas, en Vitoria en 1999, en sendos paneles realizados en el IX Congreso CLAD (Madrid noviembre 2004) (coord. J.L.Osuna), y en el X Congreso del CLAD (Santiago de Chile, octubre 2005) (coord.J.Hintze); y en el curso de verano organizado por las Cajas de Ahorro Vasco-Navarras, en San Sebastián, en septiembre de 2005.

Joan Subirats Humet
Instituto de Gobierno y Políticas Públicas
Universidad Autónoma de Barcelona

ÍNDICE

1. Los instrumentos de evaluación como palanca de gobierno en el sector público
2. Las administraciones públicas no son como las empresas
3. Tareas de gestión y tareas de gobierno
4. Tipos de relación y de poder según las distintas formas de interacción Administración-ciudadanos
5. Legitimidad de las administraciones públicas y mecanismos de rendición de cuentas
6. A diferentes tipos de organización, distintos tipos de control estructural
7. Las políticas sociales requieren procesos de evaluación específicos
8. Las políticas complejas requieren evaluaciones complejas
9. El proceso formal de evaluación tiene fuertes elementos de innovación
10. Las políticas de inclusión social no son estandarizables *a priori*
11. La falta de adecuación entre evaluación y diseño de nuevas políticas
12. La época actual exige políticas heterogéneas, flexibles y próximas
13. En el futuro será más complicado evaluar políticas sociales
14. El trabajo en red es hoy determinante en las políticas de inclusión social

Referencias bibliográficas

N.º de clasificación JEL: H43

Palabras clave: políticas sociales, procesos de evaluación

1. LOS INSTRUMENTOS DE EVALUACIÓN COMO PALANCA DE GOBIERNO EN EL SECTOR PÚBLICO

Parecería absurdo plantearse una pregunta de esta índole salvo, simplemente, para enfatizar retóricamente una afirmación. La respuesta lógica y razonable, y que compartimos, es: «sin duda». Y es la respuesta lógica y razonable porque los parámetros dominantes en las labores de análisis crítico del funcionamiento de las administraciones públicas hacen hincapié, precisamente en la falta de incorporación de instrumentos de evaluación en sus prácticas habituales de gestión y gobierno, lo que dificulta, se dice, la capacidad para mejorar la actuación de esas mismas administraciones. Lo curioso es que, tanto la «doctrina» académica como las aportaciones del mundo de la consultoría y la propia reflexión interna de los gestores públicos parecen

estar de acuerdo en la necesidad-bondad de incorporar instrumentos de evaluación en las labores de gobierno y gestión de las políticas y las administraciones públicas, pero se avanza muy poco en el tema. Una vez más, tenemos una notable brecha entre consenso teórico y práctica político-administrativa. En este texto se pretende profundizar en esa brecha y más precisamente, en los motivos que explicarían ese desajuste entre teoría y práctica. Se propondrán también algunas vías de avance.

2. LAS ADMINISTRACIONES PÚBLICAS NO SON COMO LAS EMPRESAS

Desde nuestro punto de vista, uno de los problemas centrales con que nos enfrentamos en las estrategias dominantes de modernización de las administraciones públicas es que siguen partiendo del paradigma de que Administración y empresa se parecen. Hay una parte de verdad en esto. Si en ambos casos existen objetivos, tienen que manejar recursos, dirigir y ser dirigidas por personas, trabajan mediante procesos, etc., ¿por qué no acercar métodos de funcionamiento, organización y control?. Pero a pesar de todas las similitudes las diferencias prevalecen. Después de unos cuantos años de predominio de ese modelo de cambio en las administraciones públicas, y tras constatar sus positivos impactos en la mentalidad y en la filosofía de actuación de algunas de ellas (no tan significativamente en otras, como el caso de las españolas[2]), es hora de darnos cuenta, también, de las diferencias. Como afirma Metcalfe (1993), ha llegado el momento en que ya no basta imitar, es necesario innovar.

Uno de los elementos centrales de la estrategia empresarial es la capacidad de competir. Una empresa competitiva es una empresa eficaz y eficiente. En las administraciones públicas, la lógica de actuación se basa más en la capacidad de trabajar de forma coordinada y en estrecha interdependencia que en la capacidad de competir. Son excepción las actuaciones administrativas que puedan desarrollarse con plena independencia de otros servicios u otras administraciones públicas. Los procedimientos formales o informales de cooperación y los procedimientos explícitos o implícitos de coordinación son determinantes para asegurar una gestión eficaz. Como bien sabemos, es precisamente a partir de la constatación de las continuas interdependencias en la gestión pública de dónde surgió toda la literatura que puso de relieve las limitaciones de la teoría de la decisión racional en su aplicación al contexto de las administraciones públicas (Lindblom, Wildavsky, Landau, *et al*). Las administraciones públicas trabajan en campos y en sectores que les vienen, en muchos, casos predeterminados normativamente. No escogen los mejores «nichos» de actividad. Actúan, muchas veces, de forma contradictoria, ya que sirven a intereses y clientelas también contradictorias. Su labor, lejos del ideal que expresa el concepto de «intereses generales» (Crozier 1995), se enmarca en un entramado de gran complejidad, lleno de actores, algunos de ellos muy potentes, que manejan intereses con-

[2] Para ver un balance del proceso de cambio en las administraciones públicas españolas, ver: Gallego,R. –Subirats,J., 1998, «¿Podremos sostenernos en Europa con una Administración sin renovar?», en Informe España 1997, Fundación Encuentro, pp. 269-334

trapuestos. En ese contexto, las administraciones públicas no pueden jugar con la relativa facilidad que ofrece el hecho de tener una orientación inequívoca hacia una única cuenta de resultados basada en la obtención de beneficios. Si en la empresa mercantil, eficacia y eficiencia tienden a confundirse, no siempre ocurre así en el mundo de las administraciones públicas. Equivocarse en ello significa, muchas veces, contribuir a los bien conocidos fenómenos de la llamada «selección adversa», de la cual tenemos buenos ejemplos en sanidad y educación.

A pesar de todo, sabemos que no podemos admitir que exista una insuperable colisión entre los principios de eficacia y legalidad, y por tanto hemos de trabajar de manera innovadora para ir construyendo categorías que nos permitan poder integrar los valores que expresan ambos requerimientos, a la vez que aseguramos siempre la capacidad de rendir cuentas, la capacidad de mantener abiertos los controles que eviten los abusos de poder. Una organización (aquí sí, pública o no pública) que mantenga su capacidad de rendir cuentas, de ser transparente en su gestión, de evaluar su quehacer, será una organización capaz de aprender tanto de sus aciertos como de sus errores.

3. TAREAS DE GESTIÓN Y TAREAS DE GOBIERNO

Ahora bien, al margen de estos primeros elementos de reflexión, al introducir mecanismos de evaluación y de rendición de cuentas en el funcionamiento de las administraciones públicas, uno de los elementos centrales que deberíamos tener en cuenta es saber si estamos tratando de evaluar temas que son propios del área de gobierno, o temas que son propios del área de gestión. Sabemos perfectamente que no hay nada más difícil que distinguir con precisión ambas esferas, pero también sabemos que si confundimos absolutamente esos campos se nos complicará muchísimo cualquier vía de mejora e innovación. Proponemos que se distinga, al menos analíticamente, entre lo que llamaríamos tareas de gestión y tareas de gobierno, y que se crucen con los distintos ámbitos de actuación de las administraciones, unos de contenido operacional, otros de contenido estratégico (ver cuadro n.º 1).

Cuadro n.º 1

Tareas y ámbitos de actuación de las Administraciones Públicas

	Operacional	Estratégico
Gestión	1	2
Gobierno	3	4

Fuente: Elaboración propia a partir de Metcalfe, L., *Accountability and effectiveness*, Venice, 1998, mimeo

Sumario

El cruce entre tareas de gestión, de carácter más interno, y ámbito operacional (casilla 1), podríamos decir que genera el espacio típico de los análisis preocupados por los valores de economía y eficiencia. Es el espacio en el que opera la clásica labor de intervención en muchas administraciones de nuestro entorno, y es también el ámbito en el que con mayor intensidad se han producido las reformas que se han nucleado en torno al concepto de *New Public Management (NPM)*. El espacio de gestión de contenido más estratégico (casilla 2), sería aquel en que las variables que cuentan son las de eficacia y gestión del cambio. Usando los conceptos puestos en circulación por Osborne y Gaebler, no se trataría sólo de «remar» (casilla 1) sino de «timonear» (casilla 2). Las labores emblemáticas serían las de adaptar las operaciones a los problemas, previendo posibles desajustes de la gestión. Sería este un campo del que, en nuestro país, se han estado ocupando gestores con capacidad directiva probada, o bien políticos que han actuando y actúan más en el campo de la gestión que en el propio de la política y del gobierno.

El entorno de tareas de gobierno es un terreno en el que se lidia, sobre todo, con el exterior de la organización, dónde se juegan las grandes cuestiones de la legitimación política y social. Si esa labor se cruza con el ámbito más estrictamente operacional (casilla 3), nos encontraremos con lo que podríamos denominar labores de «contramaestre», gobierno de operaciones, y con funciones como contraste, supervisión, control del gasto y de los recursos, precauciones para evitar fraudes o corruptelas, etc. Es el terreno donde solemos encontrar los órganos de control de legalidad presupuestaria y contable. El terreno menos explorado es aquel donde se entrecruzan gobierno y ámbito estratégico (casilla 4). Siguiendo la parábola marítima que hemos venido utilizando, sería el campo propio de la navegación, de las labores de «capitanía». Es aquí dónde debería determinarse lo importante y significativo para cada organización. Es ahí donde se juega la legitimidad de los objetivos de las administraciones públicas. Si no se llena ese espacio, faltan criterios de consistencia para poder emprender una labor evaluadora seria. Si no existe una carta de navegación, ni remar ni timonear tiene demasiado sentido, y el papel de los controladores pierde significación.

4. TIPOS DE RELACIÓN Y DE PODER SEGÚN LAS DISTINTAS FORMAS DE INTERACCIÓN ADMINISTRACIÓN-CIUDADANOS

Siguiendo esa línea de reflexión, podemos suponer que sería posible ir más allá del campo estrictamente operacional en el que solía refugiarse el NPM. Los mecanismos de control, de rendición de cuentas, podrían contribuir a ir llenando de contenido ese componente estratégico de la gobernación de los asuntos y la gestión pública. Muchas de los interrogantes que el NPM plantea (¿es bueno o no separar contratador y proveedor?, ¿es mejor crear agencias o mantener sistemas centralizados de provisión de servicios?, ¿ventajas e inconvenientes del uso de contratos e indicadores?, ¿es posible insistir en eficiencia sin erosionar la equidad o provocar fenómenos de selección adversa?, ¿deben generalizarse los mecanismos de copago?, etc.), podrían en-

Sumario

Cuadro n.º 2

Tipos de relación y tipos de poder en liza en distintas formas de interacción Administración-ciudadanos

Sujeto	Usuario	Cliente	Ciudadano
Administrador Jerarquía	Proveedor Intercambio	Profesional Expertise	Representación Influencia Legitimación

Fuente: Elaboración propia a partir de Metcalfe, L., *Accountability and effectiveness*, Venice, 1998, mimeo

contrar vías de respuesta más integrada si se plantearan en un contexto no sólo operacional, sino también estratégico y de gobernación general del sistema. Y para ello puede ser muy importante distinguir entre distintos tipos de organización, diferentes tipos de tareas y distintos tipos de usuario.

Por mucho que las tendencias encuadradas en la «nueva gestión pública» hayan generado efectos positivos al usar el término «cliente» como contrapunto radical a las tradicionales definiciones de «administrado» o «súbdito», lo cierto es que la generalización de ese concepto de cliente y lo que lleva aparejado, sin tener en cuenta las distintas relaciones y estructuras de poder que conllevan los distintos marcos de actuación de las administraciones públicas, puede acabar produciendo más confusión y efectos contraproducentes que resultados positivos.

Las diferentes administraciones públicas intervienen en distintos campos y prestan multitud de servicios diversos. Esas interacciones generan relaciones entre las partes que se acercarían a lo que podríamos calificar como diversos tipos ideales (ver Cuadro n° 2).

En unos casos predomina la clásica relación Administración-sujeto (típica de una Administración básicamente reguladora, inspectora y sancionadora); en otros casos, esa relación se acerca a la que se da en la esfera mercantil entre usuario y proveedor (cuando la Administración simplemente suple o actúa como prestadora de servicios o bienes que también podrían tener como protagonistas a empresas y usuarios); encontramos, asimismo, la relación característica que se da entre clientes y profesionales (esencial en servicios como sanidad y educación); y, finalmente, detectamos la que más se acerca al mundo de la política y la democracia representativa: la que se da entre los ciudadanos en general y sus representantes políticos (cuando estos últimos canalizan intereses o valores).

Podemos afirmar que en cada una de esas interacciones surgen relaciones de poder o de influencia que tienen características específicas. Jerarquía, intercambio mercantil, *expertise*, o influencia-legi-

timación, se dan en cada caso de manera predominante. Deberemos ir con sumo cuidado para evitar aplicar de manera indiferenciada unos mismos parámetros de control o unos mismos mecanismos de rendición de cuentas a interrelaciones basadas en lógicas muy distintas; o simplemente acabar confundiendo esas distintas relaciones y esas distintas posiciones de poder o de influencia con indicadores no pensados específicamente para recoger los distintos valores que se encierran en cada situación. Si, por ejemplo, en una universidad (dónde podríamos encontrar los cuatro tipos de relación y de poder mencionados) primamos de manera casi exclusiva el control y la evaluación de resultados desde una lógica mercantil o de intercambio, y marginamos o pasamos por alto en la práctica otros valores que se esconden en otros tipos de relación también presentes (las propias de la corporación de profesores o las específicas de la universidad como institución representativa y ciudadana), las tensiones y distorsiones que provocaremos serán monumentales.

No podemos, pues, simplificar ni tratar de reducir la complejidad presente en la actuación de la mayoría de administraciones públicas, usando como único parámetro de evaluación y de mecanismos de rendición de cuentas el que encierra la lógica de intercambio, competencia y eficiencia (que tan buenos resultados ha dado en el mundo de la iniciativa mercantil, dónde la complejidad ha sido hasta ahora mucho menor).[3] Es absurdo buscar relaciones de cliente donde no las hay, cómo es absurdo tratar de construir sistemas de competencia dando alternativas de salida (*exit*) a los ciudadanos, cuando quizás lo que se busca es hacer más operativa e influyente la opción voz (*voice)* usando las conocidas expresiones de Hirschman, (1970).

5. LEGITIMIDAD DE LAS ADMINISTRACIONES PÚBLICAS Y MECANISMOS DE RENDICIÓN DE CUENTAS

Si lo que nos interesa es examinar la eficacia de las organizaciones públicas, deberíamos poner atención al tipo y la calidad de las interrelaciones entre esas organizaciones públicas y el exterior, ya que si sólo nos dedicamos a examinar los componentes internos del funcionamiento de tales organizaciones, lo máximo que conseguiremos será avanzar en la evaluación de su eficiencia y del grado de cumplimiento de las procedimientos previstos.

Debemos ser conscientes de que no será posible mantener (sin sufrir erosión) la legitimidad de las administraciones públicas si no existen mecanismos potentes de rendición de cuentas. La ausencias de estos o su relevancia deteriora la base misma de existencia de las administraciones públicas. Las administraciones públicas ostentan poder, tienen capacidad de imponer su voluntad, siempre y cuando mantengan la capacidad de rendir cuen-

[3] Lo cierto es que también en el mundo de las empresas privadas se da una creciente importancia a cuestiones como los balances ecológicos y los balances sociales, como forma de complementar y aumentar el valor añadido de unas empresas que tendían sólo a ser vistas como productoras de plusvalía. Pero, a pesar de ello, continúa siendo cierto que será la cuenta de resultados la que acabará determinando, finalmente, la buena o mala marcha de la entidad.

tas, de ser rendidores de cuentas (*accountables)* frente a la ciudadanía, de seguir siendo «legítimas».

Hace unas décadas, las administraciones públicas estaban sometidas al escrutinio público desde unos parámetros que respondían esencialmente a la voluntad de evitar extralimitaciones, desviaciones o abusos de poder. El objetivo se centraba en impedir o reducir las posibilidades de tales excesos. Todo aquello que implicara apartarse del camino previsto, de lo definido previamente, era interpretado como una puerta abierta a posibles extralimitaciones. En estos momentos, junto con el mantenimiento de la exigencia de evitar esos posibles abusos, existe una clara preocupación (expresada en España por el artículo 9.2 de la Constitución y por su habilitación intervencionista[4]) por lograr que las administraciones públicas sean eficaces, consigan resultados, trabajen para construir políticas que favorezcan la plena igualdad y libertad. El problema es que, si bien en el primero de los supuestos (el más clásicamente liberal de evitar extralimitaciones en la acción pública) la estructuración y operacionalización de los mecanismos de control se ha hecho sistemática y concienzudamente, no ha ocurrido lo mismo en el segundo de los campos mencionado (el que deriva de una lógica más socializadora). Y si no existen o no se han construido criterios claros de eficacia que recojan ambos supuestos, lo que ocurre es que existe información sesgada (estrictamente sobre la legalidad en la actuación administrativa) y desorientación respecto a si está siguiendo o no el camino constitucionalmente previsto, y no se reciben señales claras y coherentes sobre los logros políticos y sociales obtenidos.

6. A DIFERENTES TIPOS DE ORGANIZACIÓN, DISTINTOS TIPOS DE CONTROL ESTRUCTURAL

Diferentes tipos de organización, con diferentes tipos de relación y de poder en su seno, deberían generar distintos tipos de control estructural (ver cuadro n.º 3). Por ejemplo, en el caso de las organizaciones basadas sobre todo en las reglas del intercambio y de la influencia, se dará mucha más importancia a la capacidad de responder a las demandas externas a la propia organización, que cuando se trate de otro tipo de organizaciones más orientadas hacia el interior (las basadas, por ejemplo, en la *expertise* o incluso en la jerarquía).

Por tanto, desde nuestro punto de vista, no se trata tanto de ofrecer pautas generales y universales de rendición de cuentas y de evaluación, sino más bien de entender concretamente y diferenciar las distintas perspectivas que pueden existir entre organizaciones, o dentro de una misma organización, para segmentar y establecer parámetros adecuados de evaluación y de control.

Hasta ahora, en las administraciones públicas se ha restringido mucho esa lógica diferenciadora que, a nuestro entender, sería capaz de revelar muchas más dimensiones de valor. Ha primado siem-

[4] «Corresponde a los poderes públicos promover las condiciones para que la libertad y la igualdad del individuo y de los grupos en los que se integra sean reales y efectivas; remover los obstáculos que impiden o dificultan su plenitud y facilitar la participación de todos los ciudadanos en la vida política, económica, cultural y social». Art.9.2 de la Constitución española, 1979

Cuadro n.º 3

Tipos de organización, tipos de poder y tipos de control

	Autoridad jerárquica	Intercambio	Autoridad basada en *expertise*	Influencia
Valores y normas	Ejercicio de responsabilidad administrativa	Interés propio, ganancias mutuas	Compromiso profesional con el cliente	Consenso, solidaridad y participación
Controles estructurales	Apelación a autoridades externas	Competencia entre proveedores	Evaluación del rendimiento basada en los «pares»	Control democrático de los representantes

Fuente: Elaboración propia a partir de Metcalfe, L., *Accountability and effectiveness*, Venice, 1998, mimeo

pre la lógica más fundacional, la clásica visión liberal en la que primaba, por encima de todo, el cumplimiento de la legalidad. Debemos recordar que esa visión partía de una profunda desconfianza de la actuación de los poderes públicos, sobre los que se extendía la poderosa sospecha de que tendían a extralimitarse constantemente en el ejercicio de sus poderes. No pretendemos decir que esa perspectiva no sea importante. Tratamos, simplemente, de señalar que existen otras perspectivas igualmente significativas, y que la propia (nueva) legalidad constitucional lo recoge (como hemos visto en el caso de España en nuestra referencia al artículo 9.2 de la Constitución).

En este sentido, los Tribunales de Cuentas, que han ejercido hasta ahora una labor positiva desde una perspectiva analítica determinada, deberían trabajar (con los cambios institucionales y de personal necesarios) para ofrecer a los organismos representativos de los que dependen unos instrumentos de control y evaluación que permitieran disponer de parámetros fiables para cada tipo específico de organización y de prestación; o alternativamente su trabajo debería complementarse con otros organismos o mecanismos que permitieran recoger la medición de valor obtenida por la labor de las administraciones públicas, que no puede reducirse, simplemente, al cumplimiento de legalidad. No darse cuenta de esto sólo produciría confusión y frustración, ya que a los organismos se les exigirían unos resultados de la labor desempeñada que no están en condiciones de dar.

La riqueza y pluralidad de perspectivas y valores que hemos presentado exige, desde nuestro punto de vista, tratar de responder a la pregunta que encabeza este texto desde una perspectiva plural. La res-

puesta sigue siendo positiva. Los instrumentos de evaluación pueden ser usados como palancas de gobierno, con la condición de que esos instrumentos de evaluación sepan recoger la riqueza de matices y de perspectivas que toda acción de gobierno genera. Una de esas perspectivas es la legal. La que examina si la acción gubernamental se ha desarrollado según lo estipulado por las previsiones legales. No obstante, deberíamos ser capaces de responder a otras preguntas. ¿Podría haberlo hecho de manera más eficiente y eficaz una organización no pública?. ¿Se perderían con ello valores que no se recogen en un simple análisis coste-beneficio?. Pero, asimismo, podríamos preguntarnos si la labor desempeñada responde a los parámetros corporativos y profesionales que están en la base de ese servicio o prestación. La calidad del servicio se calibra, en este caso, no sólo según las previsiones normativas o a la relación coste-beneficio, sino según aquellos indicadores que el conjunto de profesionales de ese ámbito entiende como básicos para poder referirse al servicio como «de calidad».

Pero las administraciones públicas democraticas no deben olvidar que su poder está condicionado, es indirecto. Su capacidad de actuación, su propia razón de ser, deriva de que logren mantener su legitimidad, su influencia sobre el conjunto de la ciudadania, lo que significa que la ciudadanía siga entendiendo que esas administraciones son «sus» administraciones. Y, por esta causa, los valores que le otorgan esa capacidad de influencia son (muchas veces) distintos de los que influyen para decir que una organización es eficiente o competitiva. El hecho de que la actuación de unas administraciones sea solidaria con aquellos sectores o ámbitos territoriales que no disponen de recursos suficientes para alcanzar una calidad de vida digna, el hecho de que «los individuos y los grupos en que se integran» (art. 9.2 de la Constitución) puedan participar de manera efectiva en el diseño y puesta en práctica de las decisiones y políticas públicas, acaban, sin duda, influyendo en la percepción sobre la legitimidad de la actuación de unos poderes públicos que han situado a la participación democrática entre uno de sus valores fundamentales. Lo mismo sucede con la sensación de que las decisiones del Ejecutivo responden, mayoritariamente, a la voluntad de la ciudadanía (y no totalmente al contrario, como ocurrió en el caso de la decisión del gobierno del PP de involucrar a España en la invasión de Irak). Todos estos parámetros influyen indudablemente en la evaluación de la actuación de los poderes públicos. Es forzoso reconocer que quedan lejos de parámetros estrictamente de eficiencia o de cumplimiento de la legalidad.

Es, por tanto, poco «eficiente» tratar de reducir el debate sobre la evaluación de las administraciones públicas a un mero ejercicio técnico sobre cuáles son los indicadores oportunos o cuál debería ser el órgano técnicamente más adecuado para realizar esa labor. Cuando intentamos evaluar la actuación pública no podemos reducir la complejidad de la actuación de los poderes públicos, y su amplia gama de criterios sobre los que basar la calidad (y «calidez») de su actuación, a un problema de traslación de las concepciones y dinámicas que han funcionado, y en parte siguen funcionando, en el ámbito mercantil. Hemos de trabajar en ese entorno y encarar la complejidad y pluralizar métodos y protagonismos.

Sumario

En definitiva, el debate de hoy sobre el papel y los resultados de la acción pública ya no solo se sitúa en el *cómo* hacer las cosas, sino que exige contar con instrumentos que permitan responder previamente a preguntas tales como: *de qué* deben ocuparse los poderes públicos y *quién* es el más indicado para ofrecer resultados eficaces y eficientes a esas demandas sociales canalizadas por dichos poderes representativos.

7. LAS POLÍTICAS SOCIALES REQUIEREN PROCESOS DE EVALUACIÓN ESPECÍFICOS

¿Todo lo que hemos dicho hasta aquí sirve para cualquier política pública?. ¿Sirve para las políticas sociales?. Cada política pública tiene características específicas, y un buen proceso de evaluación debería tener en cuenta esas especificidades. Pero, en el caso de las políticas sociales, esto es especialmente significativo. No creo que sea necesario dedicar un gran espacio a argumentar y justificar las especiales dificultades que genera la evaluación de políticas sociales y, más en concretamente, la evaluación de políticas de lucha contra la pobreza y a favor de la inclusión social.

En esta parte del artículo trataremos de explorar los enfoques que permitan evaluar el grado de consecución en los objetivos de inclusión social de las políticas que se han puesto en marcha en España y Europa. Como sabemos,[5] la exclusión social tiene componentes de multidimensionalidad y de trayectoria personal y colectiva muy específicas, que aumentan la complejidad implícita en toda evaluación de resultados y de consecución de valor por parte de la actuación de los poderes públicos y otros agentes. La puesta en marcha de planes de inclusión social desde 2001, en toda Europa, puede permitir un primer balance crítico.

Como ya anticipábamos, las políticas sociales acostumbran a ser políticas mucho más complicadas de evaluar que otras. En primer lugar porque, normalmente, tienen perfiles poco concretos. Es decir, la definición de objetivos suele ser relativamente abstracta y ambigua ya que, por otro lado, este tipo de abstracción y falta de concreción en la definición de objetivos favorece la obtención de consenso para estas políticas. Por lo tanto, no se trata de vituperar o menospreciar la forma en que estas políticas definen sus objetivos, sino de comprender el marco cognitivo, el entramado institucional y social que rodea su formulación, para entender su complejidad y las dificultades que puede generar su evaluación desde las distintas perspectivas de las instituciones implicadas, los beneficiarios reales y potenciales, y el resto de actores implicados en la política o actuación concreta.

Además, la propia Unión Europea (UE), en su afán de buscar no tanto el medio común denominador de las políticas que se llevan a cabo en los distintos países miembros, sino el modo de incentivar procesos de innovación en un campo tan complejo

[5] Ver, al respecto, el volumen que recoge el trabajo colectivo del IGOP sobre el tema, «Pobreza y exclusión social. Un análisis de la realidad española y europea», Joan Subirats (dir.), Estudios Sociales n.16, Fundación La Caixa, Barcelona, 2004, disponible totalmente en PDF en www.estudios.lacaixa.es, estudios sociales nº 16, y la bibliografía que lo acompaña

como el de las políticas de inclusión, ha introducido nuevos factores de complejidad en la evaluación de tales iniciativas. En efecto, no es fácil encontrar parámetros relativamente estandarizados, ensayados y calibrados anteriormente, para evaluar estas políticas donde predomina un proceso de «coordinación abierta» que quiere ser respetuoso con la diversidad de situaciones de partida de cada país en materia social (desde la tradición sueca hasta la mediterránea, por poner un ejemplo), y con las grandes diferencias de prioridades y situación de exclusión que existen entre los diferentes países.[6]

Pongamos un ejemplo derivado de las políticas de desarrollo económico local que se llevaron a cabo hace unos años en la UE, basadas en la nueva idea de los «yacimientos de empleo». Con esto se hacía hincapié en la necesidad de encontrar nuevos filones de ocupación que hasta entonces hubieran sido poco explotadas. Ese nuevo concepto requería ser comprendido por parte de los distintos actores institucionales, los agentes sociales, los consultores y los gestores públicos. Y, lógicamente, se necesitaba más tiempo para construir indicadores que lograran distinguir un «yacimiento» real de lo que es simplemente una nueva forma de subvencionar a los desempleados. Por tanto, no existe una total automaticidad entre la innovación en políticas y la capacidad de evaluar esas mismas políticas, ya que toda política realmente innovadora implica cambios en los elementos conceptuales, sustantivos y operativos con que debe desarrollarse y, evidentemente, evaluarse.

8. LAS POLÍTICAS COMPLEJAS REQUIEREN EVALUACIONES COMPLEJAS

Por otro lado, la evaluación de políticas de inclusión choca con otros elementos más generales. Como sabemos, a la propia dificultad en definir este tipo de políticas, cabe añadir la multiplicidad de objetivos que acostumbran a rodear este tipo de decisiones. Objetivos múltiples (mejorar la formación académica, las condiciones de salud, la integración familiar, las redes relacionales, etc.), a veces contradictorios (por ejemplo, incrementar el nivel formativo sin dejar fuera a nadie; conseguir empleo para todos al tiempo que se mejora la competitividad), y por tanto difícilmente evaluables con una misma gama de parámetros. Por otra parte, muchas veces las instituciones financiadoras están muy imbuidas de una lógica económica que tiende a considerar como unívocos conceptos como «coste» o «beneficio». Y, como sabemos, coste y beneficio son conceptos que están sujetos (o deberían estarlo) a un intenso debate social y político desde las distintas perspectivas, situaciones y posición respecto a la distribución de recursos.

También es importante señalar la contradicción existente entre los motivos que impulsaron la creación de la política (motivos que tienden a ser complejos y fruto de un proceso en absoluto lineal), y la necesidad de contar con indicadores sencillos que rápidamente den respuesta a los resultados y los impactos conseguidos (a costa, muchas veces, de reducir

[6] Sólo recientemente se habla de empezar a «sincronizar» las políticas sociales de los distintos países miembros.

Sumario

las complejidades inherentes a la formulación de la política sin que exista, además, una tradición consolidada sobre cómo evaluar este tipo de nuevas políticas). Es decir, podemos tener políticas innovadoras, relativamente complejas tanto en su formulación como en su articulación (con diversidad y multiplicidad de objetivos), e indicadores y procesos de evaluación relativamente pobres, muy condicionados por dinámicas de justificación del gasto y de explicitación de resultados, y sujetos, además, a calendarios políticos y administrativos que no siempre coinciden con el ciclo de vida y desarrollo de las políticas.

Por todo ello, es frecuente que los indicadores acostumbren a ser relativamente insuficientes o excesivamente simples con relación a la complejidad de la política emprendida, cosa que ya nos indica la necesidad de combinar metodologías *hard* y *soft* en el diseño de los procesos de evaluación de las políticas de inclusión social, ya que si utilizamos sólo aproximaciones cuantitativas conseguiremos un resultado que tenderá a ser indudablemente más pobre que el que obtendríamos asumiendo el principio que dicta: «en políticas complejas, evaluaciones complejas».

9. EL PROCESO FORMAL DE EVALUACIÓN TIENE FUERTES ELEMENTOS DE INNOVACIÓN

Por otro lado, debemos ser conscientes de que la visión que se tiene desde el territorio «afectado» por la intervención pública, tanto de la formulación de la política como de su proceso de evaluación, suele ser relativamente escéptica. Sobre todo cuando las bases conceptuales y la formulación general de esa política se originan muy lejos del lugar en que deberá hacerse operativa. No es extraño percibir en los territorios la sensación de que más que una política pensada para resolver problemas, se dispone de «soluciones en busca de problemas». Y, como estas «soluciones» vienen acompañadas de recursos, lo normal no es rechazar esa política por extemporánea y retórica sino adaptarse «creativamente» a ella e intentar acercar lo que se ha decidido desde lejos a la realidad concreta del territorio para, de este modo, obtener los recursos que se necesitan para aquello que se considera prioritario.

Asimismo, hay que tener en cuenta que, muchas veces, en los procesos de evaluación coinciden distintos objetivos o metas. Por un lado, encontramos la necesidad por parte de los actores del territorio de responder a las exigencias que plantea la Comisión Europea u otra instancia financiadora. Es decir, deben rendirse cuentas, y hacerlo de una forma determinada; y, por otro lado, existe la necesidad de avanzar en procesos de aprendizaje en el propio territorio y de los actores implicados en relación a esas políticas (que son, en muchos casos, innovadoras), ya que plantean procesos de colaboración entre actores de distintos niveles de gobierno y de diverso estatuto formal, donde no hay costumbre ni tradición de hacerlo. Por esto, la relativa «pobreza» de ciertos resultados que llegan, por ejemplo, a Bruselas, de un proceso de evaluación correctamente hecho pero realizado para cumplir los parámetros necesarios de *accountability* que están reglamentariamente previstos, no debería confundirse con la posible riqueza de de-

sarrollo de una evaluación más real y más «útil», que muchas veces puede llegar a producirse en el territorio donde se ha llevado a cabo.

¿Dónde ponemos el énfasis, en el análisis del cumplimiento formal o en el proceso de aprendizaje? La mejor respuesta es, obviamente, tratar de cumplir con ambos requerimientos, ya que responden a valores y lógicas distintas pero igualmente pertinentes. Lo peor, no obstante, es confundir y mezclar ambas lógicas, ya que nuestra experiencia nos indica que, muchas veces, lo que es valioso desde una lógica no lo es tanto desde la otra. Pero es evidente que si el proyecto incorpora un requerimiento formal de evaluación, no se debe menospreciar ni tomárselo como una obligación sin valor añadido alguno. El proceso formal de evaluación puede y debe ser aprovechado para conseguir mecanismos de respuesta y señales concretas sobre el proceso de implementación, desviación de objetivos, relaciones gestión-resultados, aspectos, todos ellos, que pueden ser perfectamente aprovechables en secuencias posteriores.

10. LAS POLÍTICAS DE INCLUSIÓN SOCIAL NO SON ESTANDARIZABLES *A PRIORI*

A pesar de todo, existe la sensación de que la evaluación que incorpore el análisis de los elementos procesuales acabará teniendo mejores impactos en los procesos de mejora e implementación posteriores. Las políticas de inclusión social tienen, casi siempre, fuertes elementos de innovación, de adaptación creativa a situaciones heterogéneas y no estandarizables a priori y, precisamente por eso, los elementos de proceso son muy significativos. No debería, pues, extrañarnos que las evaluaciones que tienen en cuenta estos procesos singulares puedan llegar a ser notablemente más potentes y útiles que las que se centran exclusivamente en el tema «resultados».

Tampoco podemos dejar de señalar la aparente contradicción que existe entre las «obligaciones» formales que rodean los procesos de evaluación de políticas sociales vinculadas a dinámicas europeas, y su utilidad práctica. En efecto, muchas veces los propios gestores de los proyectos asumen la evaluación de estos como una carga burocrática más que les imponen los financiadores, sin que las contrapartidas sean evidentes. La experiencia tiende a demostrar que si la evaluación se realiza con seriedad, y con los requerimientos profesionales necesarios, trasciende el mero cumplimiento de la formalidad prevista y «retorna» valor a los decisores y gestores del proyecto, además de contribuir a definir mejor los objetivos en el futuro y facilitar, desde un principio, la previsión de los indicadores que será necesario generar e introducir. En un campo como el de las políticas sociales, esto no es precisamente irrelevante dada la poca costumbre que hay de pensar en términos de indicadores de eficacia al plantear las acciones que se van a desarrollar. Por lo tanto, a pesar de la crítica que pueda hacerse de los aspectos más formales, lo cierto es que la mera incorporación de esa «obligación» en contextos administrativos poco acostumbrados a la cultura de la rendición de cuentas es, en sí, claramente significativa porque obliga a pensar en objetivos y resultados, cuando

lo que predomina es la mera comparación entre recursos disponibles y recursos utilizados, es decir la lógica de control contable de los *inputs*.

Al mismo tiempo, en la lógica propia de las políticas sociales será importante no desdeñar el tema de la cultura de gestión que ha predominado en esa política sectorial. En los temas sociales predomina una visión muy de «caso», centrada en la vicisitud personal concreta con la que se enfrenta el gestor público. No es fácil salir de esa «cultura» (con grandes ventajas de personalización y adecuación específica a las características de la persona que uno tiene delante), para tratar de contemplar los problemas desde una perspectiva más genérica que permita, por ende, construir categorías analíticas susceptibles de ser evaluadas una vez operacionalizadas y desarrolladas las políticas que de ellas se deriven. Tenemos, por tanto, otro «resultado» positivo derivado de la introducción de una perspectiva de evaluación de políticas de inclusión: favorecer perspectivas de trabajo que operen con mejores relaciones entre los casos específicos y las categorías con las que suelen operar las políticas.

11. LA FALTA DE ADECUACIÓN ENTRE EVALUACIÓN Y DISEÑO DE NUEVAS POLÍTICAS

El tema de la utilidad, al que hemos aludido varias veces, es importante. Las preguntas relevantes, desde mi punto de vista, serían: ¿hasta qué punto se aprende en los procesos de evaluación?, y ¿hasta que punto este aprendizaje está luego realmente presente en las nuevas decisiones o políticas que deben impulsarse? Y, en este tema, el fallo no es achacable a las políticas sociales de manera específica, sino que constituye un problema general del modo en que habitualmente opera la UE, simplemente porque los procesos de evaluación no están bien situados en la dinámica decisional de dicha institución. Lo tradicional es que, cuando el proceso de evaluación ha concluido, casi siempre el nuevo programa de carácter plurianual que debería haber sido diseñado realmente sobre la base de esa evaluación ya esté en marcha debido a que las dinámicas políticas, presupuestarias y administrativas de la UE así lo exigen. Con lo cual, la sensación general es que las evaluaciones acaban teniendo poca repercusión en los procesos decisionales. Esto no significa que esas evaluaciones no tengan efecto alguno, o que sean inútiles o simplemente retóricas; no obstante, desde el estricto punto de vista del calendario, cuando se entrega la evaluación ya está en marcha el nuevo «*call for projects*» de la nueva oleada en que se concretan las decisiones de las distintas Direcciones Generales con relación a las distintas políticas, cosa que, en apariencia, debería hacernos dudar de la real efectividad de lo que hemos denominado como procesos de aprendizaje.

Lo cierto es que los gestores de la UE y los supervisores generales de la marcha de los proyectos y planes de acción reciben información de las evaluaciones sin esperar a que estas concluyan. Existen informes a medio proyecto, reuniones de intercambio, etc. Pero esa realidad menos formalizada debería exteriorizarse más para evitar la sensación general de desencaje entre dinámi-

cas de evaluación y dinámicas de decisión. Algo podría lograrse si, en los propios protocolos de decisión, se diera valor explícito a las «lecciones» aprendidas en anteriores fases. Al mismo tiempo, habría que poner especial cuidado en no acentuar los aspectos más burocráticos de la evaluación, para no dar la impresión de que no se valoran ni se tienen en cuenta aspectos importantes en las evaluaciones de políticas sociales, que son más difícilmente cuantificables, menos tangibles, menos susceptibles de ser operacionalizados por medio de indicadores. En caso contrario, se acentúa exclusivamente la necesidad de que las «plantillas» se rellenen y los números cuadren para poder recibir la nueva remesa de recursos con los que esperamos seguir avanzando y mejorar aquellos aspectos aprendidos que, curiosamente, son los que quizás menos han servido para completar los formularios de evaluación.

12. LA ÉPOCA ACTUAL EXIGE POLÍTICAS HETEROGÉNEAS, FLEXIBLES Y PRÓXIMAS

Desde un punto de vista más general, estamos atravesando un proceso que, desde el punto de vista de un observador y analista de las políticas sociales, complica enormemente los procesos de evaluación. Como es sabido, nos hallamos en un escenario en el cual las tradicionales políticas sociales típicas del Estado de bienestar que se consolidó en Europa entre los años 1945 y 1980, tienen graves problemas de adaptación a las nuevas realidades de exclusión, desigualdad y necesidad de respuesta a los nuevos escenarios económicos, sociales y familiares provocados por la globalización. Las políticas convencionales de bienestar funcionaban, básicamente, a partir de cuatro elementos característicos: homogeneidad, rigidez, indiferencia y lejanía. Esto permitía desarrollar políticas relativamente universales en grandes ámbitos territoriales. En cambio, la época en que vivimos exige, cada vez más, políticas que sean heterogéneas, flexibles, tendencialmente personalizadas y que puedan desarrollarse desde la proximidad. Esos cuatro nuevos elementos, que juzgamos imprescindibles para dar respuesta a las nuevas necesidades sociales, obligan a que los mecanismos de evaluación recojan esa nueva realidad y esas nuevas exigencias. En cambio, la propia estructuración de la UE, su misma configuración tiende, lamentablemente, a mantener la lógica de lejanía, una cierta indiferencia ante realidades diversificadas, y una cierta rigidez en los procesos.

En este sentido, es significativa la propuesta surgida de la propia UE que plantea que son, precisamente, los procesos de gobernanza, de trabajo en redes, de articulación de los distintos niveles de gobierno, los que permitirían evitar la aparición de una creciente brecha o separación entre las dinámicas y necesidades de cada territorio y las dinámicas comunitarias, político-administrativas, que generan las políticas, los proyectos y, en definitiva, la disposición de fondos para dar respuesta a los nuevos y renovados problemas de inclusión social. La complejidad de la exclusión social no puede abordarse como si abordáramos estrictamente los temas de pobreza o de desempleo, cosa que implicaría una evidente simplificación. Como sabemos, si hablamos de exclusión social hablamos

de multifactorialidad, de multidimensionalidad, de necesidad de implicar a la comunidad en los procesos de respuesta, de participación de los actores y, en muchos casos, nos referimos a trayectorias vitales muy específicas y no de dinámicas indiferenciadas. Cuando, por ejemplo, la UE detecta con acierto la significación de un potencial de empleo entre los que nunca han trabajado anteriormente (los llamados «inactivos»), la puesta en práctica del proyecto supone buscar esas personas, detectar esos colectivos. Y, cada vez más, esa necesidad de dar respuesta a dinámicas que son menos estandarizables será una exigencia que nos planteará la propia realidad. Será a partir de esta realidad que deberemos buscar y articular respuestas desde el espacio público (no confundir con los poderes o instituciones públicas, estrictamente), respuestas colectivas a problemas colectivos en ese espacio público concreto. En tal caso, la lógica de actuación de la UE debería ser, como en parte se ha podido intuir a lo largo de este texto, mucho más la de habilitar o posibilitar que la de decidir, ya que ese proceso de habilitación es el que lograría dar respuestas a esa situación de heterogeneidad, de personalización, de cercanía y de flexibilidad a la que antes se hacía referencia.

13. EN EL FUTURO SERÁ MÁS COMPLICADO EVALUAR POLÍTICAS SOCIALES

La impresión que predomina, ya que los políticos sociales serán más complejas, con más actores. Por otra parte el proceso será, generalmente, mucho más importante que los resultados concretos, con muchos más casos personales y específicos que genéricos y estandarizables, y en un contexto en el que la transversalidad y la integridad tendrán un relieve mucho mayor que la sectorialidad (cuando en cambio estamos muy acostumbrados a trabajar sectorialmente, departamentalmente, «ministerialmente»). Pero, de seguirse una lógica más descentralizadora en la implementación y la implicación social de esas políticas, puede surgir una cierta facilidad en los procesos de evaluación al dar más valor y «voz» a la proximidad respecto a la comunidad en que se llevan a cabo esas políticas y acercar, así, los procesos de evaluación al lugar en el que se está realmente trabajando en ese campo, sin aceptar, lógicamente, que cualquier medio vale, pero estableciendo un diálogo mucho más amplio con el territorio. De esa indudable complicación puede derivarse, por contra, más capacidad de aprendizaje.

Esto exigirá, probablemente, un cambio en la lógica. Abandonar las posiciones de centralidad de Bruselas, entendidas como posiciones de neutralidad y de estar por encima de los «intereses sectoriales y territoriales», y acercarse a las situaciones reales, no para controlar o inspeccionar lo que ocurre sino para ofrecer recursos, ayuda, para reforzar la capacidad de transformación e impacto.

14. EL TRABAJO EN RED ES HOY DETERMINANTE EN LAS POLÍTICAS DE INCLUSIÓN SOCIAL

Finalizaremos con algunos comentarios sobre como avanzar en la evaluación de

los mecanismos de red, de gestión de redes, de gobernanza cuando, precisamente en las políticas de inclusión social, se pone crecientemente de relieve la significación de la implicación colectiva y de las lógicas de colaboración público-no público en sus procesos de formulación e implementacion. Éste es, sin duda, un déficit significativo en estos momentos. ¿Pueden evaluarse este tipo de cuestiones? ¿Puede evaluarse, por ejemplo, la confianza interorganizativa? Asumamos que una red viene marcada por la densidad de relaciones de interdependencia de sus nodos, de continuidad en esa interdependencia, y por la no-existencia de un centro jerárquico en esa red. Si esos son los tres elementos claves, ¿cómo podemos evaluar la capacidad de gobernanza de esa red?, ¿su capacidad de respuesta colectiva a unas demandas a las cuales la red intenta dar respuesta de una manera interdependiente, de una manera continuada y sin lógicas jerárquicas?

Podemos tratar, en principio, de analizar los procesos de formación de la red. Es decir, trabajar con lo que podríamos denominar el cálculo de costes y beneficios que hacen los actores que componen esa red. Los diversos actores tienen sus propias expectativas respecto a los procesos de colaboración que se van a emprender. Tienen sus propios puntos de vista sobre los problemas que deben resolverse. Y tienen, o no, experiencias previas de colaboración con el resto de actores de la red. Por otro lado, convendría examinar el grado de consolidación de los procesos de confianza mutua entre los actores de la red. ¿Hasta qué punto conocen, comprenden y comparten los objetivos del resto de participantes? De esa comprensión dependerá la construcción de vínculos de confianza y de metas comunes. La manera de construir indicadores al respecto dependerá de la exploración de criterios cualitativos, de procesos de cooperación entre las distintas fases del proyecto. Encontrar parámetros que nos permitan hablar de avances y retrocesos en la capacidad de generar «capital social» basado en la confianza, reciprocidad, cooperación e intereses comunes es un reto.

Por otro lado, la confianza no solo se desarrolla y se consolida cuando se comprenden los objetivos de los demás, sino también cuando se pueden compartir total o parcialmente esos objetivos, entender las formas de actuar del resto de actores y, de alguna manera, compartir esas formas de actuación. Es decir, deberíamos poder hablar de compatibilidad cultural y de compatibilidad estratégica ya que, evidentemente, es el punto clave para la consolidación de esa red porque estamos intentando construir mecanismos de respuesta colectiva a problemas colectivos, como los que acaparan los esfuerzos de las políticas de inclusión. Se trata, por tanto, de conseguir evaluar la gestión de las interacciones mediante el análisis de cómo se ha procedido a la formación de la red, cuáles han sido los criterios de inclusión, qué reglas de negociación se han seguido, qué se ha hecho para dinamizar esa red, qué espacios concretos interinstitucionales se han creado para dar respuesta a la heterogeneidad de los actores, y qué mecanismos de resolución de conflictos interorganizativos se han estructurado para dar respuesta a los problemas a medida que surgían. Si aceptamos que el trabajo compartido, el trabajo en red, es hoy de-

terminante en las políticas de inclusión, debemos reconocer que la búsqueda de estrategias para tener en cuenta el valor generado por la construcción de esa red, y la evaluación de su eficacia, su capacidad de aprendizaje conjunto y su fuerza para el futuro, son elementos cada vez más significativos en toda dinámica de evaluación de las políticas sociales actuales.

Sumario

REFERENCIAS BIBLIOGRÁFICAS

CROZIER, M. (1995): «La posición del Estado ante los otros actores», *Gestión y Análisis de Políticas Públicas,* número 2, pp. 93-100

HIRSCHMAN, A.O. (1970): *Exit, Voice and Loyalty,* Ed. Harvard University Press, Cambridge (Ma).

METCALFE, L (1993): «Public Management: From Imitation to Innovation», en Kooiman (ed.), *Modern Governance,* SAGE, pp.173-190.

Sumario

Evaluación y gestión eficiente en el sector público: aspectos organizativos e institucionales

El artículo detalla los aspectos organizativos e institucionales que influyen en el desarrollo de una gestión pública eficiente y en el funcionamiento de los sistemas de evaluación de actividades públicas. Desde un enfoque de la Economía de la Organización, se analizan aquellos procesos que permiten una gestión eficiente de la información y que son la base del funcionamiento adecuado de los sistemas de control interno y externo. Asimismo, se revisan los requisitos básicos exigibles para la implantación de sistemas de incentivos basados en resultados, instrumento esencial en la consecución de una gestión pública eficiente. La elección óptima de los instrumentos empleados para desarrollar la gestión pública es también analizada desde una perspectiva organizativa. El artículo concluye con una comparación entre dos modelos de arquitectura organizativa, uno descentralizado y con autonomía en la gestión operativa, y otro centralizado y con jerarquía en la adopción de decisiones de gestión.

Artikuluan kudeaketa publiko eraginkorraren garapenean eta jarduera publikoen ebaluazio-sistemen funtzionamenduan eragina duten antolaketaren eta erakundearen alderdiak azaldu dira. Antolaketaren Ekonomiaren ikuspegitik, informazioaren kudeaketa eraginkorra ahalbidetzen duten eta barne eta kanpoko kontrol-sistemen funtzionamendu egokiaren oinarria diren prozesuak aztertu egin dira. Era berean, emaitzetan oinarritutako pizgarri-sistemak ezartzeko eska daitezkeen oinarrizko baldintzak berrikusi dira, pizgarria kudeaketa publiko eraginkorra erdiesteko funtsezko bitartekoa den aldetik. Kudeaketa publikoa gauzatzeko erabiltzen diren bitartekoen hautaketa zuzena ere antolaketaren ikuspuntutik aztertu da. Artikuluari bukaera emateko bi antolamolde-ereduren arteko konparaketa egin da: bat deszentralizatua eta eragiketen kudeaketaren gainean autonomia duena, eta bestea zentralizatua eta kudeaketa-arloko erabakiak hartzeko hierarkiaz baliatzen dena.

The present article gives details about the organisational and institutional aspects that have influence on the development of an efficient public management and on the function of the evaluative systems of public activities. These processes, that allow an efficient management of information, are analysed from an Organisation Economy focus, and which are at the same time the bases of the adequate function of the systems of intern and extern control. Likewise, the basic demanded requirements are checked for the establishment of the incentive systems that are based on results, which is an essential instrument at the securing an efficient public management. The optimum election of the instruments used in the development of the public management is also analysed from an organisational perspective. The article concludes on a comparison of two types of an organisational architecture; a decentralized one with an autonomy in the operational management, and another one which is centralised and has an hierarchy when adopting management decisions.

Jorge Onrubia Fernández
Universidad Complutense de Madrid

ÍNDICE

1. Introducción
2. Eficiencia, organización y gestión en el sector público
3. El valor de la información en los procesos de asignación presupuestaria
4. La motivación individual en las organizaciones públicas
5. La elección óptima de los instrumentos de gestión
6. Alternativas de arquitectura organizativa para el sector público
7. Consideraciones finales
Referencias bibliográficas

Palabras clave: gestión pública, eficiencia organizativa, incentivos, información, evaluación de políticas públicas

N.º de clasificación JEL: D29, D73.

1. **INTRODUCCIÓN**

Desde la década de los años ochenta del siglo XX, el debate acerca del comportamiento más o menos eficiente de las organizaciones que integran el sector público ha ocupado un lugar importante tanto en la agenda de los gobiernos como, sobre todo, en los foros de discusión académicos. Con frecuencia, si bien en unos países más que en otros, la discusión ha aparecido mezclada con apreciaciones ideológicas sobre el papel que debería desempeñar el sector público en las modernas economías de mercado. Asimismo, el interés por estos temas reflejaba, en gran medida, un cambio de actitud de la sociedad frente a la intervención pública, especialmente en lo que se refería a sus actividades provisoras, con un cuestionamiento de su eficacia. No debe pasarse por alto, además, la influencia que ha podido tener el saneamiento de las finanzas públicas al que se han debido enfrentar la mayor parte de los países desarrollados. Detrás de los desequilibrios macroeconómicos y fiscales aparecían, en no pocas ocasiones, políticas públicas muy generosas en sus prestaciones y poco estimulantes con el esfuerzo productivo individual, lo que a la larga condicionaba el mantenimiento del propio estado del bienestar.

Esta aproximación inicial a la cuestión de la eficiencia en el sector público orientó las propuestas y medidas reformadoras en una línea clara: la reducción del peso del sector público en la economía, con un desplazamiento hacia el sector privado de muchas de aquellas activida-

des hasta aquel momento desempeñadas desde los gobiernos. La política decidida de aligeramiento de la presencia del sector público en Estados Unidos y en el Reino Unido caló en casi todos los países desarrollados, si bien con procesos muy diferentes. En el caso de España, al igual que en otros Estados miembros de la Unión Europea como Francia o Italia, el proceso fue mucho más lento, aunque iba a verse precipitado por la necesidad de sanear las cuentas públicas para el acceso a la última fase de la Unión Económica y Monetaria.

El fenómeno de la globalización ha traído nuevos retos para los sectores públicos, que se ven presionados a ofrecer soluciones eficaces a importantes problemas asignativos, como la necesidad de potenciar la libre competencia, la promoción de los procesos de I+D+i, la supresión de las distorsiones generadas por los sistemas fiscales, la mitigación de los problemas de deslocalización de factores productivos e inversiones, la acumulación de capital humano y tecnológico o la dotación de infraestructuras. Tampoco los retos en materia de redistribución y lucha contra la pobreza son despreciables, destacando la lucha contra el desempleo estructural generado por los cambios tecnológicos y de localización, la integración de la inmigración, incluidas la lucha contra la marginalidad, y la extensión de las prestaciones básicas de las coberturas sociales. En este escenario, la necesidad de dar respuestas satisfactorias tanto a las imperfecciones del sistema de mercado como a los problemas de desigualdad y pobreza existentes emerge como un objetivo primordial de cualquier sociedad que desee alcanzar cuotas elevadas de bienestar social e individual.

La experiencia acumulada en más de dos décadas permite comprobar que, a pesar de los intensos procesos de privatización y liberalización habidos en los países más desarrollados de la OCDE, la obtención de estas mejoras continuadas de bienestar no puede conseguirse apelando únicamente al libre funcionamiento de los mercados. Lejos de ello, cada vez parece más imprescindible su adecuada combinación con intervenciones públicas que promuevan intercambios voluntarios más eficientes, y faciliten para toda la población niveles aceptables de consumo de bienes privados y públicos. Sin embargo, la oportunidad, idoneidad y eficacia de las actuaciones del sector público no son propiedades intrínsecas que vengan garantizadas por el carácter público de la acción colectiva. No es pequeña, tampoco, la experiencia que informa de la falta de efectividad de muchos programas públicos, su excesivo coste o las importantes distorsiones de toda índole generadas en su aplicación.

En definitiva, y sin cuestionar la importancia que para el adecuado funcionamiento de la economía ha traído la racionalización de los programas presupuestarios, no puede desconocerse la trascendencia que para el crecimiento económico y el bienestar social tiene contar con organizaciones públicas eficientes. Si la provisión de determinados bienes y servicios individuales y colectivos o las actuaciones de redistribución de renta no pueden ser resueltas por el libre funcionamiento del mercado, y si la mera titularidad pública de las organizaciones encargadas de estas actividades no garantiza respuestas eficaces a las necesidades que están detrás de todas esas actividades, nos situamos ante un trascendental dile-

Sumario

ma que no podemos demorarnos en resolver. A mi juicio, su resolución solamente puede alcanzarse desplazando el análisis desde los argumentos que justifican la intervención pública hacia la consideración de la gestión pública como un proceso de selección de las alternativas más eficientes y, por tanto, como un problema organizativo inherente a cualquier proceso de asignación de recursos económicos.

El denominado «enfoque positivo» de la Economía Pública ha ofrecido, tradicionalmente, una explicación al inadecuado funcionamiento del sector público basada, principalmente, en los comportamientos estratégicos de legisladores, burócratas y grupos de interés. Aunque esta teoría de «fallos del sector público» ha facilitado la comprensión de un buen número de las limitaciones y deficiencias a las que se enfrentan las organizaciones públicas en la búsqueda de respuestas eficaces a los problemas apuntados, no es menos cierto que esta aproximación ha sido bastante parca a la hora de ofrecer alternativas que permitan superar satisfactoriamente los fallos asignativos y distributivos del mercado[1].

Un nuevo enfoque de análisis que explica los problemas de ineficacia del sector público como una consecuencia de un comportamiento organizativo ineficiente se ha ido abriendo paso en los últimos tiempos. Esta aproximación parte de considerar al Estado como una organización económica compleja, encargada de la asignación y la distribución de recursos, además de la regulación de las condiciones de acceso a los mercados (Stiglitz, 1989). En consecuencia, las respuestas a encontrar son, esencialmente, soluciones de gestión eficiente, y deben ser buscadas en el ámbito del análisis microeconómico, donde juega un papel destacado la moderna economía de la organización.

La naturaleza del problema de la eficiencia del sector público es esencialmente microeconómica, por más que sus consecuencias se proyecten inevitablemente sobre la política fiscal (González-Páramo y Onrubia, 2003). Por tanto, cualquier iniciativa de reforma encaminada a mejorar la eficiencia del sector público tiene que ser abordada con un proceso de diseño organizativo-institucional orientado a la búsqueda de un arquitectura capaz de hacer compatibles los aspectos micro y macroeconómicos de la gestión pública[2]. Cualquier instrumento, mecanismo o proceso que sea necesario incorporar a la consecución del objetivo de una gestión pública más eficiente debe ser considerado como un elemento definitorio de esa solución o respuesta organizativa. Como razonaremos más adelante, la incorporación de sistemas de evaluación, externos o dentro de la función de control interno de la organización, no debe ser ajena a esta exigencia, pues, en definitiva, se trata de uno de los elementos clave a la hora de aportar la información y generar los incentivos indispensables para promover comportamientos organizativos eficientes.

---

[1] Sobre las potencialidades y limitaciones de la «teoría de fallos del sector público» para explicar el comportamiento del sector público, y su conjunción con el enfoque normativo de la «teoría de fallos de mercado», generalmente empleada para argumentar el porqué de las intervenciones públicas, puede verse el trabajo de Wolf (1993).

[2] En Onrubia (1995) se revisan los fundamentos de la relación entre los aspectos micro y macroeconómicos de la gestión pública.

La reflexión sobre estos condicionantes organizativos e institucionales ocupa el contenido de este artículo, cuya estructura es la siguiente. La segunda sección presenta, a la luz de un enfoque de economía de las organizaciones, las premisas básicas de la relación entre organización y gestión eficiente en el ámbito del sector público. La tercera sección aborda la importancia que el valor de la información tiene para conseguir un comportamiento organizativo eficiente, prestando especial atención al papel que debe desempeñar en los mecanismos de asignación presupuestaria. La sección cuarta analiza otro elemento fundamental del diseño organizativo en el sector público como es la motivación individual. En concreto, en esta sección se revisan los principales condicionantes a los que se enfrentan las organizaciones públicas a la hora de incorporar mecanismos de incentivos al desempeño. La relación entre eficiencia en la gestión y selección óptima de los instrumentos empleados en el desarrollo de la misma es tratada, de forma breve, en la sección quinta del artículo. El análisis se cierra, en la sección sexta, con una comparación entre los dos modelos tipo de arquitectura organizativa habitualmente empleados en los países desarrollados para la realización de las actividades encargadas al sector público. El artículo concluye con una síntesis valorativa.

2. EFICIENCIA, ORGANIZACIÓN Y GESTIÓN EN EL SECTOR PÚBLICO.

En los últimos años, se ha abierto paso un nuevo enfoque para tratar de comprender los factores que están detrás del comportamiento eficiente del sector público. Esta aproximación parte del reconocimiento de la relevancia que poseen los aspectos institucionales y organizativos a la hora de evaluar el comportamiento más o menos eficiente que muestra el sector público en el desarrollo de sus actividades. Esta aproximación entronca con los contenidos más tradicionales de la Economía de las Organizaciones, especialmente con aquellos desarrollados a partir de la Economía de la Información, de la teoría de los costes de transacción y de la teoría de los contratos incompletos[3]. En el origen de este enfoque encontramos a autores como Stiglitz (1989) o Tirole (1994), quienes identifican en los problemas de actuación del sector público una ausencia de respuestas adecuadas a los problemas de coordinación y de incentivos a los que éste se enfrenta como cualquier organización de naturaleza económica[4].

Por consiguiente, la búsqueda de una mayor eficacia del sector público debe si-

[3] En un reciente trabajo, Gibbons (2005) ofrece una interesante revisión sobre el estado actual de estos diferentes enfoques de la Economía de las Organizaciones y sus consideraciones particulares a la hora de analizar la empresa como arquitectura básica de organización. Muchas de sus reflexiones, en busca de una síntesis formalizable, son de aplicación al análisis de las organizaciones públicas.

[4] El trabajo de Stiglitz (1989) constituye fundamentalmente un ensayo acerca de las potencialidades y limitaciones del sector público en la corrección de fallos de mercado y actividades de redistribución de renta. El trabajo es, en cierto modo, un intento de redefinición de la teoría de fallos de mercado, con el propósito de ofrecer un catálogo de intervenciones factibles y no el tradicional catálogo de diagnósticos. Toda la discusión se centra en las capacidades y limitaciones del sector público como organización, si bien se trata de un trabajo de reflexión no formalizado, aunque con importantes vías de desarrollo. En Onrubia (1999) se ofrece una extensión de su contenido para extraer propuestas de gestión pública, a la luz de los principales resultados disponibles.

Sumario

tuar a los responsables públicos, tanto técnicos como políticos, ante la resolución de problemas de eficiencia organizativa, que son, en definitiva, los que determinan el desarrollo de la gestión. Como argumenta Stiglitz (1989, 1994), estos problemas son, en muchos casos, inherentes a la propia noción de organización y, por tanto, independientes de su titularidad pública o privada. No obstante, aunque el análisis parta del reconocimiento de este carácter organizativo, dichos problemas no deben ser contemplados al margen de las limitaciones de naturaleza política que el Estado, por sus propias características, impone a su organización. En concreto, la naturaleza de los objetivos perseguidos y su origen electoral, los mecanismos de elección de los gestores encargados de dirigir las actividades, la utilización de incentivos de «baja potencia», el principio de subsidiariedad en el manejo de fondos públicos o las reglas particulares de exigencia de responsabilidades, condicionan de forma notable la selección de los instrumentos de coordinación e incentivos que pueden emplearse en el diseño organizativo. Además, las propias características del ejercicio de autoridad pública, como el poder exclusivo de coacción fiscal, el asociacionismo universal de sus miembros o el acceso electoral a la autoridad, que sirven para explicar las potencialidades del sector público frente al mercado, en otras ocasiones aparecen como serios condicionantes a la hora de buscar respuestas organizativas eficientes[5].

Para comprender la ruptura que significa, respecto de la tradición, esta nueva perspectiva organizativa de las cuestiones relacionadas con la eficiencia del sector público, es necesario contemplar el tratamiento que el análisis económico ha venido dando al tema. Durante los años ochenta, la enorme mayoría de los trabajos que se realizaron se ocuparon, casi en exclusiva, de la presentación de técnicas de medición de la eficiencia productiva con la que se desarrollaban distintas actividades por parte de las unidades gestoras. Inicialmente, estos trabajos cuantitativos se centraron en el ámbito de la salud y de la educación, extendiéndose progresivamente a otras provisiones públicas, como la justicia, la producción de servicios públicos locales o la gestión de impuestos. Puede decirse, a la vista de la literatura disponible, que la aproximación científica predominante en relación con la eficiencia del sector público, tanto desde un plano metodológico como empírico, es de índole cuantitativa, si bien es cierto que en los últimos años han proliferado aquellos trabajos que, a partir de los resultados alcanzados, incorporan recomendaciones de gestión dirigidas a incrementar la eficiencia productiva[6].

Posiblemente, la construcción de indicadores de gestión para distintas actividades del sector público y su inclusión dentro del diseño de sistemas de seguimiento y evaluación de objetivos constituye el área donde la vinculación entre las evaluaciones de eficiencia y la reforma de la gestión pública ha sido más fructífera[7]. La incorporación por las instituciones

[5] Estas limitaciones pueden revisarse en González-Páramo y Onrubia (2003).

[6] Knox Lovell y Muñiz (2003) ofrecen un panorama sobre el estado actual de la investigación sobre medición de eficiencia productiva con aplicación al sector público.

[7] Smith y Goddard (2003) ofrecen una interesante revisión del estado actual de esta línea de trabajo.

públicas de nuevos instrumentos y técnicas tanto en la esfera del control interno como externo ha corrido desigual suerte según de qué países, incluso de qué niveles de gobierno, se trate. Existe en la actualidad un consenso, cada vez más amplio, de que detrás de muchos intentos fracasados de implantación de sistemas de evaluación de la eficiencia emergen importantes carencias organizativas que impiden una incorporación efectiva de los mismos a la arquitectura institucional del sector público. De hecho, cuando nos enfrentamos a la cuestión de la eficiencia del sector público asumiendo el enfoque organizativo al que acabamos de hacer referencia, se está aceptando la relevancia de una serie de exigencias que debe cumplir la organización encargada de desarrollar la actividad pública, entendida ésta en toda su extensión. Sin su observancia, cualquier medida reformadora dirigida a mejorar la eficiencia en la gestión nacerá limitada en su eficacia, pues carecerá de los adecuados incentivos que han de promover su correcto funcionamiento. No debemos olvidar que la gestión, tanto en el ámbito privado como público, consiste en adoptar decisiones asignativas en ambientes con importantes imperfecciones informativas (incertidumbre, racionalidad limitada, asimetrías de información) y donde, por supuesto, todos los agentes que intervienen no tienen los mismos objetivos individuales, ni incluso colectivos.

La historia de los procesos de reforma de la gestión pública en los países desarrollados está plagada de experiencias frustradas, a las que sin duda no ha sido ajena España. Defendemos aquí la idea de que, en buena parte, estos fracasos pueden explicarse razonablemente por la ausencia de consideración organizativa de la gestión pública y, por ende, de sus problemas de eficiencia. Aunque en el ámbito académico pueda resultar obvio, conviene no olvidar que la resolución de estos problemas de índole organizativa no consiste en adoptar determinados organigramas formales, calificar nominal y jurídicamente las distintas unidades gestoras, o proponer modificaciones sin más de los catálogos de puestos de trabajo. Sin embargo, es frecuente encontrar este tipo de actuaciones «modernizadoras» en bastantes procesos de reforma, a la vez que brillan por su ausencia la adopción de mecanismos de incentivos al desempeño o vinculados a mejoras mesurables de eficiencia y efectividad, la definición precisa de áreas de responsabilidad, o el establecimiento de verdaderos contratos presupuestarios con asignaciones de fondos relacionadas con la consecución de objetivos. Incluso la realización de evaluaciones del grado de eficiencia productiva con el que se desarrollan determinadas actividades públicas no redundará en una mejora efectiva de la gestión pública si no van acompañadas de un encaje organizativo adecuado, donde queden definidos con precisión los destinatarios de los informes resultantes, las consecuencias presupuestarias, retributivas o incluso de carrera profesional que se derivan de los resultados evaluados, así como la regularidad en su realización. Únicamente así se pueden generar los incentivos necesarios para promover decisiones consistentes con lo que entendemos como gestión eficiente, es decir, aquella tendente a maximizar el valor de las actividades realizadas.

Desde los postulados de la Economía de la Información, la búsqueda de un mo-

Sumario

delo organizativo eficiente debe ser entendida como la elección de una forma concreta de encauzar las relaciones de todos los agentes que, con sus decisiones, intervienen en la realización de una determinada actividad, de manera que se mitiguen los comportamientos oportunistas y las disfunciones que generan los conflictos de objetivos, las asimetrías informativas, la incertidumbre o la racionalidad limitada[8]. En consecuencia, hacer máximo el valor de una organización exige contar con una arquitectura organizativa capaz de resolver los múltiples problemas de coordinación y de incentivos, tanto personales como institucionales, que surgen a lo largo de sus procesos de actuación y a los que, para el caso que nos ocupa, el sector público, nos referiremos a continuación.

No debemos pasar por alto, como señalan Milgrom y Roberts (1992), que aunque los aspectos jurídicos de la organización tienen su importancia, una consideración completa de la arquitectura organizativa implica considerar muchos elementos más, tales como las características de los flujos de información y de recursos, las relaciones de autoridad y de control, la distribución del poder efectivo, la asignación de responsabilidades y de las capacidades de decisión, los mecanismos para captar y conservar a sus miembros y recursos, los medios de generar y difundir nuevas ideas y conocimiento, la definición precisa de sus objetivos y fines, y los medios empleados para hacer compatibles éstos con el comportamiento de sus miembros individuales. Si nos centramos en el ámbito del sector público, el reconocimiento de la naturaleza organizativa del problema de la eficiencia en la gestión pública implica aceptar una serie de rasgos que no pueden obviarse a la hora de buscar respuestas: el carácter personal de los intereses de los agentes participantes, la particular distribución del poder de decisión, la existencia de ambientes de decisión con importantes imperfecciones informativas originadas por la incertidumbre, la racionalidad limitada, la distribución asimétrica de la información y, por supuesto, los nada despreciables costes de uso de la información inherentes a cualquier organización compleja. Por consiguiente, la consecución de niveles de eficiencia aceptables en el comportamiento del sector público, obliga a mitigar, a través de mecanismos organizativos, tanto los problemas de coordinación como los comportamientos estratégicos que, favorecidos por los conflictos de objetivos y las asimetrías de información existentes entre sus miembros, están presentes en el seno de las organizaciones públicas encargadas de la acción colectiva.

En definitiva, la búsqueda de respuestas adecuadas a los problemas de eficiencia en el sector público nos sitúa ante el reto de encontrar diseños organizativos favorecedores de una gestión eficiente.

[8] Williamson (1999) ofrece un enfoque alternativo del diseño eficiente de organizaciones, basado en la teoría de los costes de transacción. Según este enfoque, una organización eficiente debería hacer mínimos los costes de transacción propios de esa opción asignativa. Como señala Gibbons (2005), la teoría de los costes de transacción y la economía de los contratos y la información poseen muchos elementos en común. Ambas perspectivas tradicionales pueden ser usadas de forma complementaria para comprender el problema organizativo en presencia de titularidad pública. Aghion y Tirole (1997) y Rajan y Zingales (1998) han propuesto teorías alternativas basadas en la distribución de los derechos de propiedad y el poder de decisión en las organizaciones, cuya potencialidad para comprender los problemas de eficiencia organizativa en el sector público es, a mi juicio, muy destacable.

Sin embargo, esta aproximación organizativa no suele ser la habitual en los programas de reforma de la gestión planteados desde los poderes públicos. De hecho, existe una especial predisposición por parte de los políticos y burócratas, en absoluto indiferentes ante estos cambios, a proponer reformas de la gestión pública alejadas de esta visión de eficiencia organizativa. Estas reformas suelen ser, no obstante, de dos tipos, según la concepción administrativista o gerencial del sector público que, por tradición, predomine en el país en cuestión.

El primer grupo estaría integrado por reformas de carácter esencialmente «estético», donde adquiere un papel importante el fortalecimiento de los enunciados «garantistas» que pretenden favorecer los derechos de los usuarios en el acceso y uso de los servicios públicos. Por lo general, se trata de reformas articuladas mediante simples modificaciones de los organigramas de las unidades gestoras, los cuales reflejan cambios en la adscripción orgánica de los centros gestores y de las oficinas de rango inferior. Otros rasgos característicos de este tipo de procesos reformadores suelen ser la revisión de los catálogos de puestos de trabajo, las modificaciones superficiales de la normativa reguladora de los procedimientos de gestión económico-financieros, o la implantación de cambios relacionados con la imagen corporativa de los departamentos y centros gestores afectados. Es frecuente, en las mismas, su autocalificación como «iniciativas modernizadoras» de la gestión pública.

Por su parte, aquellas reformas que promueven la implantación de sistemas de información y evaluación más o menos sofisticados —contabilidad de costes, diseño de indicadores, programas de seguimiento y evaluación de objetivos, etc.—, pero en estructuras organizativas carentes de los incentivos adecuados para su correcto funcionamiento y aprovechamiento, integrarían ese segundo grupo. A pesar de la potencialidad de estos instrumentos, la ausencia de canales adecuados para los flujos de información y, sobre todo, el desinterés político y, evidentemente, administrativo, por la exigencia de responsabilidades de gestión, suele terminar convirtiendo su aplicación en simples rutinas burocráticas, estériles para lograr cualquier mejora de eficiencia en la gestión pública.

3. EL VALOR DE LA INFORMACIÓN EN LOS PROCESOS DE ASIGNACIÓN PRESUPUESTARIA

Una vez aceptada la naturaleza organizativa del problema de eficiencia de la gestión pública, merece una atención especial el papel crucial que en toda esta cuestión juega la información. Al igual que sucede en el ámbito de las organizaciones de titularidad privada, también en el sector público el «valor de la información» es, en gran medida, determinante del «valor organizativo». La complejidad de sus actividades, generalmente con múltiples objetivos a alcanzar, con una destacable dificultad para la elección de sus destinatarios efectivos, lo mismo que para prever los costes derivados de sus efectos asignativos y distributivos, hacen más valiosa, aún si cabe, la información con la que debe funcionar y ser controlada cualquier organización pública.

La Economía de la Organización ha puesto de manifiesto la importancia que

tiene para la eficiencia organizativa el concepto de eficiencia informativa[9]. Esta noción parte del hecho, incuestionable, de que la adquisición, distribución y procesamiento de la información en organizaciones complejas supone incurrir en costes económicos, ya sean monetarios o de oportunidad, nada despreciables. La eficiencia informativa exige que cualquier organización que pretenda desarrollar un comportamiento eficiente ha de verse obligada a seleccionar, entre volúmenes enormes, sólo aquella información que resulte relevante para cada agente decisor, la cual deberá serle distribuida sin fragmentaciones y con la debida oportunidad temporal. Es obvio que si los problemas de comportamiento estratégico antes apuntados no encuentran la adecuada solución en la arquitectura de la organización, la posibilidad de extracción de rentas individuales de cualquier tipo inducirá a la no revelación oportunista de información valiosa, con los consiguientes costes en términos de eficiencia, tanto para el funcionamiento de la organización como para el control externo de sus órganos de gobierno.

El problema de la eficiencia informativa alcanza a la totalidad de las actividades encomendadas al sector público, así como a las diferentes etapas que es necesario afrontar para su desarrollo, desde la elección de las áreas prioritarias de actuación, hasta la decisión sobre la forma más idónea de llevar a cabo cada intervención. Desde esta concepción organizativa, la información constituye un factor fundamental e insustituible para un correcto funcionamiento del sistema de control interno, piedra angular de la arquitectura de cualquier organización pública o privada. Resulta, en ocasiones, paradigmático contemplar la desdibujada noción que, desde visiones jurídico-administrativas de la gestión pública, se hace de los sistemas de control interno, reduciéndolos en muchos casos a los simples procedimientos de fiscalización los procedimientos de ejecución del gasto público. Esta concepción, a mi juicio, muy negativa por lo que su adopción supone de renuncia, ha supuesto en países como España, Italia o, incluso Francia, una importante merma de la capacidad de uso eficiente de la información, con sus inevitables costes en términos de eficiencia de la gestión, incluso, paradójicamente, a través de una menor eficacia de los mecanismos tradicionales de control jurídico-administrativo[10].

El valor de la información desempeña un importantísimo papel en el ámbito de la gestión presupuestaria, entendida ésta en un sentido amplio, donde su disponibilidad y tratamiento adecuado resultan cruciales para alcanzar actuaciones eficientes por parte de las unidades gestoras. Debe tenerse en cuenta que, por su carácter de no-mercado, la gran mayoría

[9] Sobre el concepto de eficiencia informativa y sus condicionantes en las organizaciones puede consultarse Milgrom y Roberts (1992, cap. 4).

[10] Desde una concepción gerencial de la gestión pública, propia del ámbito de los países anglosajones o del norte de Europa, a veces resulta bastante optimista la eficacia que se le atribuye *a priori* a los sistemas de control de legalidad que operan sobre los procesos administrativos y de ejecución presupuestaria. Su orientación hacia la preservación de los medios puestos a disposición de los gestores termina dejando en un segundo plazo, cuando no olvidando, sus consecuencias sobre los resultados alcanzados por la gestión, lo que en ocasiones puede llegar a ser una fuente de información muy valiosa para detectar comportamientos irregulares o cuanto menos no demasiado honestos.

de las provisiones de bienes y servicios realizadas por el sector público son financiadas desde el presupuesto. En un ambiente de imperfecciones informativas, Carlsen (1991), en un marco de teoría de juegos, y Onrubia (1999), en un marco de Economía de la Información, han demostrado que, si se acepta la existencia de un *trade-off* en la función de producción de actividad pública entre el coste presupuestario y el nivel de eficiencia productiva con el que ésta se lleva a cabo, sólo se alcanzará un comportamiento organizativo eficiente si se utilizan contratos presupuestarios que vinculen las dotaciones de recursos a los resultados evaluados de la actividad financiada.

En escenarios caracterizados por un conflicto de objetivos entre la dirección política de la actuación pública y los centros gestores encargados de la provisión de las actividades y donde el nivel de eficiencia productiva realmente alcanzado es información privada en poder de los responsables de las unidades gestoras, la utilización de esquemas de aportaciones presupuestarias contingentes con el volumen —ajustado por calidad— de *output* obtenido se convierte, bajo algunas condiciones, en una solución organizativa irrenunciable para paliar el oportunismo contractual. Desde esta aproximación, los «contratos presupuestarios» basados en resultados, medidos en términos de objetivos finalistas, nunca intermedios, se presentan como una solución de *second best* para promover la eficiencia en aquellas actividades del sector público financiadas de forma significativa al margen del sistema de precios de mercado[11]. De este resultado se deducen dos interesantes consideraciones respecto de la frecuente utilización de sistemas presupuestarios que, *de facto*, terminan rigiendo el proceso de asignación de fondos públicos para la financiación de actividades de provisión pública por un criterio «incrementalista». En primer lugar, su incompatibilidad con el logro de una gestión pública eficiente. Y en segundo lugar, la anulación de las posibilidades de mejora de la eficiencia que pudieran esperarse de la implantación de cualquier sistema de evaluación de resultados de la actividad pública, en la medida que este criterio de asignación neutraliza los incentivos que, a través de la información generada por el sistema de evaluación, estaría recogiendo el mecanismo del contrato presupuestario contingente[12]. En definitiva, la generalidad con la que el incrementalismo presupuestario está presente en muchos sectores públicos, y en todos los niveles de gobierno, no hace albergar demasiadas esperanzas sobre la contribución que hacen los sistemas de asignación presupuestaria a la eficiencia del sector público.

Además, debe tenerse en cuenta que la efectividad de este tipo de contratos presupuestarios va ligada a la existencia de sistemas de control de resultados capaces de proporcionar información relevante sobre todos los aspectos que conforman el desarrollo de cada actividad. Desde el punto de vista del diseño institucional, este requisito nos sitúa ante un modelo organi-

[11] En la medida que no existe posibilidad real de evaluar externamente a los centros gestores, la eficiencia productiva, los contratos presupuestarios empleados no puede ser una solución de *first best*. No obstante, un adecuado diseño organizativo debe incentivar la creación de servicios de control interno capaces de generar información de calidad que permitan la evaluación externa de los resultados.

[12] En relación con el contenido de estos «contratos presupuestarios» puede verse González-Páramo y Onrubia (2003).

zativo de las unidades gestoras en el que la exigencia de responsabilidades de gestión y los incentivos a los gestores en función de los resultados alcanzados, deben estimular a éstos a dotar a sus unidades de servicios de control interno eficaces. Como misión fundamental, estos servicios de control interno deben encargarse de proveer los sistemas de información —indicadores de gestión, procedimientos de seguimiento de objetivos y medición de eficiencia, contabilidad analítica, etc.— requeridos por este régimen de contratos presupuestarios. Pero como se ha señalado anteriormente, la incorporación de estos sistemas de información no puede producirse sin una definición precisa de canales de información ni sin un establecimiento nítido de responsabilidades de gestión, en sus diversas áreas. En caso contrario, la ausencia de una arquitectura organizativa adecuada suele reducir su aplicación a un mero requisito burocrático, lo que los señala en el medio plazo como candidatos a su eliminación.

La utilización, indispensable, que el sistema de control externo de la actividad pública ha de hacer de la información generada por estos servicios, ofrece otra importante faceta de control interno. El carácter multiprincipal de la organización pública, junto con la ausencia de control directo basado en el ejercicio de derechos de propiedad, obligan a contar con mecanismos de supervisión que ejerzan el control de resultados en el ámbito del poder legislativo[13]. Posiblemente estamos ante una de las cuestiones de mayor complejidad en lo que al diseño institucional de la acción colectiva se refiere.

Cómo superar la identificación de las mayorías parlamentarias con los intereses del gobierno constituye un reto en la implantación de instituciones de control externo realmente efectivas en el desarrollo de su cometido. La estrecha vinculación entre el poder ejecutivo y la mayoría parlamentaria que sostiene a los gobiernos —responsables últimos de la gestión de las actividades públicas— obligan a diseñar marcos de control externo que superen los esperables problemas de colusión y de falta de credibilidad. Se trata, sin duda, del elemento de cierre fundamental del sistema de supervisión, a través del cual la exigencia de responsabilidades se ha de convertir en un incentivo a la gestión eficiente.

Sin embargo, esta pérdida de control por parte de los partidos políticos no es fácilmente asumible, máxime si tenemos en cuenta que estamos ante los agentes políticos que tienen el poder de ejercer cualquier iniciativa reformadora. Sólo algunos países anglosajones y del norte de Europa, con una tradición de gestión pública nítidamente gerencial, han conseguido contar con una articulación constitucional del control externo realmente efectiva, con niveles de independencia suficientes para estos órganos y con una neutralización aceptable de los comportamientos colusivos entre los distintos poderes del Estado[14]. En definitiva, se trata,

[13] Sobre este carácter multiprincipal de la organización pública y sus implicaciones puede verse Martimort (1996).

[14] Laffont (2000) analiza formalmente el diseño institucional de reparto de poderes entre el gobierno y el parlamento, y algunas de sus implicaciones sobre el control externo de la actividad pública, tanto mediante instrumentos de evaluación directa de las actividades como a través de los incentivos generados por los mecanismos de *check and balances* incluidos en los modelos constitucionales.

fundamentalmente, de una cuestión de correcto diseño constitucional, donde a pesar de las importantes dificultades que la separación de poderes entraña, no debe renunciarse a la búsqueda de respuestas organizativas que refuercen la eficiencia en el comportamiento del sector público, ahora contemplado desde un plano institucional superior.

4. LA MOTIVACIÓN INDIVIDUAL EN LAS ORGANIZACIONES PÚBLICAS

La motivación de los agentes individuales que desempeñan las tareas que configuran la gestión de las actividades de cualquier institución es otro elemento esencial del diseño organizativo sin cuyo adecuado tratamiento no es factible conseguir un comportamiento eficiente. En el ámbito del sector público, el problema de la motivación afecta tanto a los gestores responsables del desarrollo de las distintas actividades públicas como a los empleados de ellos dependientes.

La perspectiva histórica, junto con la visión *weberiana* de la función pública, han complicado tradicionalmente el tratamiento de los problemas de motivación en las organizaciones públicas. De hecho, en muchos países, las retribuciones de los empleados públicos mantienen una escasa o prácticamente nula relación tanto con la productividad aportada por el trabajador como con el valor de la producción en la que participan. Esta concepción también suele incluir carreras profesionales basadas en criterios de antigüedad y escalafón, aunque a menudo aparezcan encubiertas en figuras que hacen referencia al mérito. Es frecuente, además, que los procesos de selección de empleados públicos se realicen mediante procesos estandarizados y, en gran medida, centralizados. Además, es habitual que, por argumentos de neutralización de la arbitrariedad derivada de la rotación política, el acceso a un puesto en la función pública tenga un carácter vitalicio, con importantes restricciones a la reasignación funcional o territorial.

Un escenario de relaciones laborales como el que se acaba de describir revela una ausencia de incentivos indispensables para el desarrollo eficiente de la gestión de las actividades públicas. No debe obviarse, además, que en las actividades públicas resulta muy complicado identificar y verificar por separado el esfuerzo de cada individuo, así como precisar en qué medida las distintas tareas realizadas contribuyen a la producción de las actividades finalistas responsabilidad de cada área, servicio o centro gestor.

La ausencia de compromiso político fuerte dificulta muy seriamente, en las organizaciones públicas, la implantación de mecanismos de incentivos basados en la evaluación del desempeño. Como señala Dixit (2002), la falta de referencias para la comparación generadas por la ausencia de competencia, la escasa concreción de muchas actividades, o las presiones corporativas para absorber como derecho adquirido cualquier retribución inicialmente flexible, complican notablemente el diseño y establecimiento de cualquier mecanismo de incentivos. Además, a diferencia de lo que sucede en el sector privado, en el caso de las organizaciones públicas no existen incentivos adicionales generados por la posibilidad de quiebra o por la amenaza de tomas de control por parte de los mercados de capitales. Y si además se tiene en cuenta que los usua-

rios de los servicios públicos no pagan directamente una parte sustancial de su coste, nos situamos en un escenario en el que resulta realmente complicado introducir incentivos individuales o colectivos al desempeño.

Puesto que los agentes que intervienen en la gestión pública soportan sólo una parte reducida de las consecuencias de sus decisiones y, como hemos visto, sus remuneraciones guardan escasa relación con su productividad, puede decirse que ésta está regida por «incentivos de baja potencia». El objetivo del diseño organizativo, en este terreno, a pesar de todas estas limitaciones a la motivación, es encontrar mecanismos incentivadores que mejoren el comportamiento colectivo de estos agentes. Para ello, resulta recomendable remover, allí donde sea posible, los criterios de regulación laboral que resulten incompatibles con los criterios favorecedores de una gestión eficiente propuestos desde la Economía de la Organización.

Interesantes propuestas en materia de incentivos al desempeño se han mostrado exitosas en algunos países donde se han aplicado. El diseño conjunto de carreras profesionales basadas en el mérito y de esquemas retributivos vinculados a la consecución de objetivos individuales y conjuntos parecen constituir una buena alternativa para mitigar los problemas de debilidad de incentivos en presencia de riesgo moral y selección adversa[15]. Las dificultades para verificar tanto el esfuerzo directo como los resultados del desempeño pueden superarse empleando mecanismos de revelación de información (Dixit, 2002). De este modo, ante la complejidad que supone la determinación del nivel de calidad con el que se proveen los bienes y servicios por parte de las organizaciones públicas, el diseño de salarios de eficiencia basados en la reputación se puede convertir en una alternativa muy adecuada frente a los esquemas tradicionales de incentivos (Shapiro y Stiglitz, 1984).

La existencia de fuertes interrelaciones entre diversas actividades públicas, lo que incluye la interdependencia de sus resultados, puede ser tratada mediante la implantación de esquemas de compensación basados en los resultados de evaluaciones conjuntas. De este modo se reducen, de forma importante, los problemas de debilidad de incentivos que aparecen por problemas de identificación de los esfuerzos individuales ante producciones complejas. Además, en escenarios de información asimétrica tipo riesgo moral con efectos importantes sobre la calidad de la actividad, Glaeser y Schleifer (1998) proponen esquemas de incentivos basados en contratos de compromiso con niveles de calidad prefijados[16]. Para corregir los problemas de colusión que aparecen cuando varias unidades gestoras que desarrollan actividades comunes poseen poder de influencia en los niveles de prestación pública, Laffont y Martimort (2000) proponen un interesante esquema de in-

[15] En Dewatripont, Jewit y Tirole (1999) se presenta un modelo teórico que permite sostener las ventajas, desde el punto de vista de la maximización del valor organizativo, de este tipo de incentivos combinados.

[16] Hart, Shleifer y Vishny (1997) aplican estos contratos a la gestión de prisiones, aunque Dixit (2002) encuentra limitaciones de selección adversa a la hora de la elección de los responsables de gestión.

centivos capaz de promover niveles eficientes de actividad.

En relación con la incorporación institucional de sistemas de evaluación del desempeño, Osborne y Gaebler (1992) establecen una serie de condiciones que han de verificar dichos sistemas basados en la definición de baterías de indicadores tanto cuantitativos como cualitativos. Estos requisitos permiten evitar problemas tradicionales de los mecanismos de incentivos, como el conocido «efecto trinquete», según el cual, después de periodos de buen desempeño los estándares de referencia suelen ser revisados al alza (Weitzman, 1980; Freixas, Guesnerie y Tirole, 1985). La solución propuesta a este problema de compleja resolución se centra en la participación de los agentes implicados en la definición de los indicadores, en el uso de un número de indicadores razonable, en la aplicación periódica de revisiones del sistema de acuerdo con la experiencia obtenida de su funcionamiento, así como la exigencia, argumentada en la sección anterior, de contar con controladores externos plenamente independientes.

5. LA ELECCIÓN ÓPTIMA DE LOS INSTRUMENTOS DE GESTIÓN

La aproximación organizativa adoptada en relación con la eficiencia de la gestión en el sector público no debe ser ajena a la necesidad de seleccionar, con estrictos criterios de eficiencia, los instrumentos más adecuados para desarrollar las intervenciones públicas. Se trata de introducir eficiencia a partir de la elección óptima de los instrumentos de intervención pública. Esta selección eficiente exige comparar las capacidades y límites del sector público como productor de servicios públicos, frente a las potencialidades alternativas de otros modos de gestión, tanto con participación privada total como parcial. Se trata de una cuestión compleja, que debe ser resuelta mediante la comparación entre los respectivos costes de transacción.

En la definición de su «teorema de la privatización», Sappington y Stiglitz (1987) proponen un interesante catálogo de los elementos que inciden en la elección óptima entre producción pública o privada de actividades colectivas. Su análisis parte de la premisa de que la mayor parte de las dificultades organizativas a las que se enfrentan tanto el sector público como el privado son, en origen, comunes[17]. Sin embargo, las características del sistema de precios y de la autoridad pública, como mecanismos asignativos, presentan diferencias que pueden ser determinantes de la superioridad o no del *contracting out*.

Esta elección, como corrobora el funcionamiento del sector público en el mundo real, es, en buena medida, una cuestión de grado, pues los costes de transacción entre alternativas varían de forma notable. En este sentido, considerar que la intervención del sector público supone la absorción total por éste de todas las tareas y funciones necesarias para llevar a cabo un actividad provisora de bienes o servicios públicos es, simplemente, renunciar *a priori* a cualquier pretensión de alcanzar un nivel aceptable de

[17] En González-Páramo y Onrubia (2003) se revisa con amplitud este «teorema fundamental de la privatización de la producción pública».

eficiencia en la gestión. Como se ha señalado anteriormente, muchas de las características inherentes a la organización pública complican el funcionamiento correcto de los incentivos, tanto colectivos como individuales, llegando incluso a su anulación. Sin embargo, no debe impedirse que la gestión pública incorpore todos aquellos instrumentos y diseños organizativos en los que estén presentes elementos característicos del sistema de mercado y que promuevan asignaciones eficientes de recursos. Entre otros, cabe destacar la creación de escenarios de competencia, el uso de señales en los precios, la adopción de marcos de decisión descentralizada, de incentivos monetarios basados en la propiedad, etc. Las opciones en este terreno son múltiples, y deben ser observadas como elementos constitutivos del diseño organizativo de las actuaciones públicas, y no como alternativas sustitutivas de éstas[18].

6. ALTERNATIVAS DE ARQUITECTURA ORGANIZATIVA PARA EL SECTOR PÚBLICO

La necesidad de resolver correctamente las relaciones, de nivel superior, entre la arquitectura organizativa del sector público, considerado éste en sentido estricto, y el poder político ejercido tanto a través del ámbito legislativo como ejecutivo, constituye un último condicionante para lograr una gestión pública eficiente. Desde la aproximación organizativa que se ha adoptado en este trabajo, la forma en la que se resuelven estas relaciones va a determinar el grado de descentralización efectiva con el que actúan las unidades del sector público encargadas de desarrollar las distintas actividades.

En la medida que se acepta la naturaleza informativa del problema del control organizativo, estamos en condiciones de vincular esta descentralización con la capacidad que la organización pública tiene para gestionar adecuadamente los importantes volúmenes de información necesarios para obtener un comportamiento gestor eficiente (Onrubia, 1999). Las soluciones, por tanto, del diseño organizativo van a situarse en torno a dos modelos polares que enfrentan, como rasgos distintivos básicos, centralización frente a descentralización en la adopción de decisiones de gestión operativas, y jerarquía frente a autonomía en los mecanismos de supervisión empleados.

La experiencia comparada en los países desarrollados ofrece dos modelos estilizados de arquitecturas organizativas para la institución pública, diferenciadas en función del mayor o menor grado de descentralización y autonomía con el que actúan sus unidades gestoras. Así, una serie de países, entre los que destacan Suecia, Reino Unido, Canadá, Estados Unidos, Holanda, Australia o Nueva Zelanda, cuya concepción de la gestión pública es esencialmente gerencial, presentan modelos de organización institucional del sector público caracterizados por elevados niveles de descentralización efectiva de la gestión pública. La autonomía en los procesos de asignación presupuestaria, de contratación y de gestión de personal suelen ser rasgos definitorios de las organizaciones públicas en estos países. Por lo que respecta a los sistemas de control de la actividad pública, estos es-

[18] Una revisión de los principales instrumentos de este tipo y experiencias en su aplicación puede verse en López-Casasnovas (dir.) (2003).

tán concebidos desde una óptica, predominante, de los resultados alcanzados.

En el modelo opuesto, encontramos países que, como Francia, Italia, Alemania, Bélgica o España, han adoptado tradicionalmente modelos de organización institucional del sector público con una fuerte centralización y con estructuras de decisión muy jerarquizadas. En estos países el derecho administrativo suele ser la principal fuente inspiradora de la gestión pública, de corte bastante procedimental y garantista. La centralización de los procesos presupuestarios, de la gestión de recursos humanos, así como de parte de la contratación, es una nota característica propia de estos sectores públicos.

En el primer modelo de arquitectura organizativa, la articulación institucional básica suele realizarse a través de la creación de «agencias», las cuales están encargadas de la gestión operativa de los programas públicos diseñados en el ámbito político de dirección estratégica. Estas agencias se incardinan en una estructura organizativa superior, que suele estar regida por una serie de principios organizativos en los que residen muchas de las ventajas atribuidas a este modelo institucional: a) separación efectiva entre las responsabilidades de gestión estratégica, situadas en el nivel político, y las de gestión operativa, de índole técnica; b) financiación de las agencias mediante contratos presupuestarios plurianuales vinculados al cumplimiento de objetivos de actividad y revisables anualmente; c) existencia de servicios de control interno en las agencias, bajo la responsabilidad de su dirección técnica; d) cotas significativas de independencia política en el control externo, orientado fundamentalmente hacia la evaluación económica de los resultados de las actividades. Estos cuatro elementos han de operar de forma cohesionada dentro de la arquitectura organizativa del sector público, garantizando un uso eficiente de la información, condición indispensable para alcanzar un nivel aceptable de eficiencia en la gestión.

En el segundo modelo de arquitectura organizativa, las unidades gestoras se encuentran, por lo general, situadas en estructuras verticales con dependencia jerárquica directa, siendo sometidas, por lo general, a un control de medios muy centralizado. Su arquitectura organizativa cuenta con los siguientes elementos característicos: a) vinculación elevada entre las áreas estratégica y operativa de la gestión, lo que favorece la dilución de las responsabilidades políticas y técnicas; b) controles internos dependientes de unidades horizontales situadas fuera del ámbito de poder de decisión de los gestores, y con una orientación casi exclusiva hacia el control de los medios puestos a disposición de los gestores, siendo su naturaleza primordial jurídico-contable; c) débil correspondencia funcional de los centros gestores con áreas bien delimitadas de actividad finalista; d) presupuestación caracterizada por negociaciones eminentemente incrementalistas, donde la relación entre las dotaciones presupuestarias y el cumplimiento de los objetivos previstos es muy débil; e) un control externo predominantemente jurídico y con dependencia política importante, donde la implantación de evaluaciones de los resultados dentro es, por regla general, escasa y casuística.

La comparación del comportamiento del sector público en cuanto a la eficiencia de la gestión, bajo ambos modelos

básicos, permite encontrar una superioridad significativa en el modelo de gestión descentralizada o «de agencias». Este resultado puede explicarse, desde la perspectiva organizativa adoptada, como la mejor capacidad de respuesta que tienen las arquitecturas descentralizadas, en términos de uso eficiente de la información, a la hora de resolver el problema del control organizativo. Precisamente, como se ha señalado, este elemento resulta crucial para el adecuado funcionamiento de los mecanismos de coordinación e incentivos que, a través del diseño de la organización pública, deben procurar una gestión eficiente. Por el contrario, los mayores costes de uso de la información que muestra el modelo de arquitectura jerárquica o de departamentos centralizados impiden dar una respuesta eficaz al citado problema del control organizativo. Este hecho impide responder, de forma adecuada, a las múltiples situaciones de oportunismo contractual que, como hemos visto, afloran en el comportamiento de las organizaciones públicas.

7. **CONSIDERACIONES FINALES**

En el artículo que ahora concluye, se ha tratado de presentar, adoptando un punto de vista organizativo, las condiciones que determinan, a través del diseño institucional, la consecución de una gestión pública eficiente. El interés por estas cuestiones nace de la experiencia comparada en los procesos reformadores de la gestión pública, habituales en las últimas dos décadas. La realidad muestra cómo aquellas propuestas reformadoras que pasan por alto estas exigencias organizativas están llamadas a fracasar, en la medida que su implantación suele resultar, en el medio plazo, escasamente valiosa para la mayoría de los agentes individuales que intervienen en los procesos que conforman la gestión de las actividades públicas.

Desde el ámbito de la Economía de las Organización, el trabajo pasa revista a los principales elementos del diseño organizativo, destacando su influencia a la hora de incorporar medidas concretas en materia de gestión eficiente de la información, control interno y externo, e incentivos al desempeño. Como conclusión principal, cabe destacar la relevancia que, para conseguir mejoras significativas en la gestión pública, tiene la elección adecuada de una arquitectura organizativa de tipo descentralizado y con plena autonomía operativa de sus unidades gestoras.

En el trabajo se apuesta por la adopción de arquitecturas organizativas descentralizadas, en línea con los conocidos como modelos de «agencias» o «divisionales», en detrimento de los modelos «departamentales», en los que existe una fuerte centralización, al menos formal, con una marcada dependencia jerárquica de las unidades encargadas de la gestión operativa. Se defiende la importancia de los costes de uso de la información como principal factor explicativo de esta recomendación. Los menores costes informativos en las estructuras organizativas del primer tipo tienen consecuencias beneficiosas sobre la efectividad de los sistemas de control, especialmente en aquellos basados en la evaluación económica de los resultados de las actividades públicas. Destaca, principalmente, la posibilidad que brindan a la implantación de contratos presupuestarios, con todo lo que esto supone de introducción de in-

Sumario

centivos para alcanzar mejoras sustanciales de la eficiencia.

Para concluir, creo importante resaltar la importancia que en el impulso de este tipo de reformas organizativas de la gestión pública tiene la definición de la estrategia reformadora. De nuevo nos enfrentamos a problemas de conflicto de objetivos en ambiente de imperfecciones informativas, especialmente asimetrías, muy importantes. Así, los comportamientos estratégicos tanto de los gestores como de los empleados públicos, tanto a nivel individual como desde su organización como grupos de presión, constituyen un freno a la adopción de muchas iniciativas esenciales. La escasa delimitación entre las responsabilidades de índole política y técnica, agravada en muchos países, como España, donde existe una permeabilidad muy elevada entre cuerpos funcionariales y responsabilidad política, ha supuesto un importante freno a la adopción de reformas organizativas realmente eficaces.

En este sentido, la creación de instituciones independientes respecto de la capacidad de influencia y negociación de los departamentos y centros gestores que lideren estos procesos se muestra como un requisito fundamental para evitar los esperables casos de colusión entre poderes prescrito por los modelos teóricos disponibles. La consecución de los niveles de independencia necesarios debe afrontarse, principalmente, a partir de la incorporación de profesionales no directamente vinculados a los colectivos implicados en la reforma. Si bien es cierto que la obtención de información sobre las circunstancias en las que se viene desarrollando la gestión requiere la participación de gestores y empleados públicos en el proceso, la experiencia revela que ésta debe producirse de forma externa al poder de decisión que lidera el proceso reformador. No se trata, en absoluto, de una tarea sencilla, pero la expectativa de obtener importantes ganancias de bienestar a través de una gestión más eficiente de los recursos puestos a disposición de la acción colectiva puede justificar sobradamente el esfuerzo y los inevitables costes políticos que este tipo de procesos aparejan.

Sumario

REFERENCIAS BIBLIOGRÁFICAS

AGHION, P. Y J. TIROLE (1997): «Formal and real authority in organizations». *Journal of Political Economy*, 105: 1-29.

CARLSEN, F. (1991): *Public sector efficiency. A game-theoretic analysis*, Ph. D. Dissertation, Norwegian School of Economics and Business Administration, Bergen-Sandviken.

DEWATRIPONT, M., JEWIT, I. Y TIROLE, J. (1999): «The economics of career concerns. Part II: Application to missions and accountability of government agencies». *Review of Economics Studies, 66: 199-217*.

DIXIT, A. (2002): «Incentives and organizations in the public sector: An interpretative review». *Journal of Human Resources, 37: 696-727*.

FREIXAS, X., GUESNERIE, R. Y TIROLE, J. (1985): «Planning under incomplete information and the Ratchet effect». *Review of Economic Studies*, 52: 173-192.

GIBBONS, R. (2005): «Four formal(izable) theories of the firm». *Journal of Economic Behavior and Organization*, 58: 200-245.

GLAESER, E. L. Y SHLEIFER, A. (2001): «Not-for-profit Entrepreneurs». *Journal of Public Economics*, 81: 99-115.

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

HART, O., SHLEIFER, A. Y R. VISHNY (1997): «The proper scope of government: Theory and application to prisons». *Quarterly Journal of Economics*, 112: 1127-1161.

KNOX LOVELL, C. A. Y MUÑIZ, M. A. (2003): «Eficiencia y productividad en el sector público. Temas dominantes en la literatura». *Papeles de Economía Española*, 95: 47-65.

LAFFONT, J. J. (2000):*Incentives and Political Economy*. Oxford University Press, Oxford.

LAFFONT, J. J. Y MARTIMORT, D. (2000): «Mechanism design with collusion and correlation». *Econometrica*, 68: 309-342.

LÓPEZ-CASASNOVAS, G. (DIR.), PUIG-JUNOY, J., GANUZA, J. J. Y PLANAS, I. (2003): *Los nuevos instrumentos de la gestión pública*. Servicio de Estudios de La Caixa, Colección Estudios Económicos, vol. 31, Barcelona.

MARTIMORT, D. (1996): «The multiprincipal nature of government». *European Economic Review*, 40: 673-685.

MILGROM, P. Y ROBERTS, J. (1992): *Economics, Organization, and Management*. Prentice Hall, Englewood Cliffs, New Jersey.

ONRUBIA, J. (1995): «El modelo-organizativo institucional del sector público: la necesaria compatibilidad de los aspectos micro y macroeconómicos». *Economía y Gestión de las Administraciones Públicas*, vol. 7, Congreso Nacional de Economía-Consejo General de Colegios de Economistas de España, Las Palmas de Gran Canaria.

ONRUBIA, J. (1999): *La organización eficiente del sector público: un estudio desde la economía de la información y una aplicación al caso español*. Tesis Doctoral, Facultad de Ciencias Económicas y Empresariales, Universidad Complutense de Madrid.

OSBORNE, D. Y T. GAEBLER (1992): *Reinventing Government. How the Entrepreneurial Spirit is Transforming the Public Sector.* Addison-Wesley, Reading, MA.

RAJAN, R. Y L. ZINGALES (1998): «Power in a theory of the firm». *Quarterly Journal of Economics*, 113: 387-432.

SHAPIRO, C. Y SITGLITZ, J. E. (1984): «Equilibrium unemployment as a discipline device». *American Economic Review*, 74: 433-444.

SMITH, P. C. Y GODDARD, M. (2003): «Los indicadores de gestión en el sector público. Fortalezas y limitaciones». *Papeles de Economía Española*, 95: 35-46.

STIGLITZ, J. E. (1989): *The Economic Role of the State*. Basil Blackwell, Oxford.

STIGLITZ, J. E. (1994): «Replanteamiento del papel económico del Estado: bienes privados suministrados públicamente», en G. López i Casasnovas (dir.). *Análisis Económico de la Sanidad*. Barcelona: Generalitat de Catalunya, Departament de Sanitat i Seguretat Social, pp. 19-47.

WEITZMAN, M. (1980): «The Ratchet principle and performance incentives». *Bell Journal of Economics and Organization*, 11: 302-308.

WILLIAMSON, O. E. (1999): «Public and Private Bureaucracies: A Transaction Cost Economics Perspective». *Journal of Law, Economics, and Organization*, 15: 306-342.

WOLF, C. (1993): *Markets or Governments. Choosing between imperfect alternatives*. Cambridge, MA: The MIT Press.

Sumario

De la evaluación a la gestión: acortar la brecha entre eficacia y efectividad

Este artículo propone tres argumentos principales: que la esencia de la gestión clínica y sanitaria, pasa por reducir la brecha entre efectividad y eficacia, que se obtiene aquello por lo que se paga y que la gestión de las organizaciones sanitarias rebasa los ámbitos de las mismas. Parecería pues, lógico, pagar por los resultados en términos de salud y utilizar el incentivo de la competencia por comparación. Sería conveniente acabar con las letanías «más sanidad», «más educación», «más I+D+i», y esforzarse en obtener mejor sanidad, mejor educación, mejor I+D+i, aplicando lo que ya sabemos. La evaluación, de cualquier tipo, proporciona conocimiento analítico y siempre es conveniente investigar y evaluar más, pero el auténtico déficit se produce en la puesta en práctica: es necesario talento implantador para convertirlo en una mejor gestión, talento que obviamente está constreñido por unas limitaciones, cuya consideración cierra el escrito.

Artikulu honek hiru argudio nagusi proposatuko ditu: lehenengoa, kudeaketa kliniko eta osasun-kudeaketaren funtsa efikazia eta eraginkortasunaren arteko jauzia txikitzea dela, bigarrena, ordaindutakoaren arabera eskuratzen dela eta hirugarrena, osasun-erakundeen kudeaketak horien esparruak gainditzen dituela. Bada, zentzuzkoa litzateke osasun-arloko emaitzen arabera ordaintzea eta konparazio bidezko lehiakortasunaren pizgarria erabiltzea. Komenigarria litzateke «osasungintza gehiago», «hezkuntza gehiago», «I+G+B gehiago» direlakoak baztertzea, eta osasungintza hobea, hezkuntza hobea eta I+G+B hobea erdiesteko ahaleginak egitea, dagoeneko dakiguna aplikatuta. Ebaluazioak, edozein motatakoa dela, ezagutza analitikoa dakar eta gehiago ikertu eta ebaluatzea beti da komenigarria, baina benetako eskasia, berriz, ekintzak gauzatzean gertatzen da: nahitaezkoa da ezarpenak egiteko talentua izatea, talentu hori kudeaketa hobea bihurtzeko, baina jakina baditu muga batzuk. Muga horiei buruzko gogoeta, hain zuzen ere, azalduko da amaierako lerroetan.

This article proposes three main arguments: that the clinical and health management acts to reduce the gap between the efficacy and the effectiveness, and that this is achieved by paying for results, and the management of the health organisations requires an active interaction with its environment. It seems logical to pay for the results in terms of public health and to use the incentive of the yardstick competition (competence by comparison). It would be suitable to put an end to the platitudes «more public health», «more education», «more I+D+i», and to strive in obtaining better health public, better education, better «I+D+i» applying what we already know. The evaluation, of any type, produces analytic knowledge, and it is normally convenient to investigate and to evaluate, but the real deficit is produced at the practical level: It is necessary to get things done, an implementation talent in order to turn knowledge into a better management, a talent which is obviously constrained by some limitations whose consideration would close this article.

Vicente Ortún Rubio
Universitat Pompeu Fabra, Barcelona

ÍNDICE

1. Reducir la brecha entre efectividad y eficacia constituye la finalidad de cualquier gestión sanitaria y clínica
2. La actual gestión sanitaria
3. Características diferenciadoras de las organizaciones sanitarias
4. La difusa frontera entre política y gestión
5. Recapitulación
Referencias bibliográficas

Palabras clave: efectividad, eficacia, gestión clínica

N.º de clasificación JEL: L97, I1, I18, H43

1. REDUCIR LA BRECHA ENTRE EFECTIVIDAD Y EFICACIA CONSTITUYE LA FINALIDAD DE CUALQUIER GESTIÓN SANITARIA Y CLÍNICA

La última línea en cualquier *cuenta de resultados* de una organización sanitaria la constituyen las muertes aplazadas y la calidad de vida que se ha ganado. Los resultados económicos deberían tener una consideración inferior a los resultados obtenidos en la mejora de la salud, incluso en las organizaciones privadas, en las que hay unos accionistas a los que satisfacer.

La eficacia refleja el umbral de posibilidades de producción para una determinada capacidad tecnológica: es aquello que puede conseguirse si todo funciona bien. La efectividad mide el grado de consecución de ese potencial máximo en función de los datos de la realidad; la calidad de las organizaciones sanitarias no es el único dato —influyen otras organizaciones no sanitarias, la cultura poblacional, el conjunto del sistema sanitario...—, pero sí el más relevante para los gestores de servicios sanitarios. De hecho, la calidad de un proveedor puede medirse por la brecha entre efectividad y eficacia que se le puede atribuir.

Por otra parte, sería de esperar que en las organizaciones sanitarias más relevantes —públicas y privadas sin finalidad lucrativa— se intentara obtener una buena *cuenta de resultados*, recordemos: reducción de la mortalidad y mejora de la calidad de vida, sin sufrir eventuales cortocircuitos. Sabemos que hay más de una

cuenta de resultados, a veces casi tantas como integrantes de la organización. Cada uno cuida su parcela confiando en que el planteamiento global para aumentar la efectividad en el tratamiento de las enfermedades cardiovasculares, el cáncer o los accidentes corresponde a la política sanitaria. Lo que ocurre es que la *mano invisible* del mercado no funciona en este campo y la excesivamente visible mano de la política no suele ser suficiente.

Existen ciertas asociaciones básicas, atajos mentales tan cómodos como irracionales, que llevan a vincular la respuesta sanitaria con cualquier problema de salud. Por ejemplo, se asocia el hecho de que intervenga un experto en tuberculosis a cualquier aumento de la incidencia de la misma. Y eso no es así. La mayor parte del conocimiento sanitario está en las organizaciones de prestación de servicios, no en los ministerios ni en otras autoridades sanitarias. Ilustrémoslo con algunos ejemplos.

Tuberculosis

Una enfermedad conocida por el hombre desde el alba de la historia y que alcanzó proporciones epidémicas (la peste blanca) con el proceso de urbanización que acompañó a la Revolución Industrial. Antes de 1944, año en que se dispuso de un tratamiento eficaz, la mortalidad disminuyó considerablemente a raíz de las mejoras socioeconómicas, dejando al margen los colapsos del pulmón implicado, el pneumotórax terapéutico, las toracoplastias y demás intervenciones sanitarias. Los ensayos clínicos aleatorios acompañaron desde el inicio, y de manera pionera, a los nuevos tratamientos y a las medidas preventivas introducidas en la década de los cuarenta. La elevada eficacia de los mismos se puso rápidamente de manifiesto.

Décadas más tarde, la efectividad de control de la tuberculosis deja muchísimo que desear en muchas regiones del mundo. En una reunión anual del Fondo de Investigaciones Sanitarias de la Seguridad Social (Córdoba, 1988), el ponente, refiriendo datos ajenos, hablaba de un 3,8% de efectividad en la lucha contra la tuberculosis, lo que significa que solo 3,8 personas de cada 100 tuberculosos se beneficiarían del programa de control. El 0,038 resultaba de multiplicar un 0,95 de eficacia por un 0,1 de diagnósticos sobre casos nuevos por un 0,8 de tratamientos correctos sobre casos diagnosticados por un 0,5 de cumplimiento con el tratamiento. ¿De qué servía una eficacia del 0,95 o sea del 95%? Para publicar tal vez; para *cuidar la viña* no.

Desafortunadamente investigar sobre efectividad supone llamar al mal tiempo y no hacer currículo (aunque sí biografía); los hallazgos no tendrán validez universal y, salvo que se produzca una estrepitosa innovación metodológica, no merecerán ningún eco especial. Sin embargo, resulta que el conocimiento de la efectividad es lo que más debería preocupar a un profesional sanitario, a un gestor, a un ciudadano o a un político. En España, el 70% de los programas de control de la tuberculosis de las Comunidades Autónomas no cumplían los criterios básicos recomendables para conseguir esa relación de coste-efectividad que podría encabezar cualquier liguilla de intervenciones sanitarias. Y no cumplían los criterios básicos ni en porcentaje de búsqueda activa, ni en cumplimiento, ni en terapia directamente

Sumario

observada, ni en retraso diagnóstico, ni en porcentaje de contactos revisados; esto indica la precaria situación de España en cuanto al control de la tuberculosis (Rodrigo *et al*. 1999). Y nada de ello podía achacarse a que no existan profesionales competentes en materia de tuberculosis.

También se debe mencionar el caso de India, cuya investigación sobre tuberculosis era puntera y beneficiaba al mundo entero excepto a la propia India, el país con más casos de tuberculosis en el mundo y en el que menos de la mitad de los pacientes tuberculosos eran diagnosticados y menos de la mitad de estos eran tratados efectivamente. Los servicios de laboratorio estaban infrautilizados, los regímenes de tratamiento eran innecesariamente complicados, la escasez de medicamentos era frecuente y no se valoraba el cumplimiento con el tratamiento; sin embargo, más tarde empezaron a reducir la brecha entre efectividad y eficacia: diagnóstico sobre esputo, tratamiento observado directamente, estandarización de tratamientos, etc. Esto supuso: 200.000 trabajadores sanitarios entrenados, 436 millones de personas con acceso a los servicios, 3,4 millones de pacientes evaluados y 800.000 tratados con una tasa de éxito del 80%. El coste por paciente curado era de 50 dólares aplicando lo que ya se sabía, pero no se aplicaba (Khatri *et al*., 2002).

Inadecuación en la indicación y realización de cesáreas

La organización de guardias obstétricas en llamada, atención al parto a cargo de matronas —con organización por turnos, implicación del médico únicamente en caso de riesgo o complicación— y la relación continuada y exclusiva de una organización sanitaria con sus médicos, ilustra cómo la adecuación en la indicación y realización de cesáreas es susceptible de una buena gestión. Explica la neutralización tanto de la tendencia a indicar y realizar cesáreas por razones de conveniencia (descanso, ocio), como la de realizar cesáreas por otras consideraciones no estrictamente clínicas (Spetz *et al*., 2001).

Se obtiene aquello por lo que se paga

What gets measured gets managed. ¿Y qué medimos en nuestros sistemas sanitarios? En primer lugar, aquello que se presupuesta o compra; en el mejor de los supuestos, actividad. Actividad que también puede denominarse *producto intermedio* si se quiere resaltar el carácter instrumental de los servicios sanitarios como otro factor más de la producción de salud. Es sabido que este producto intermedio presenta dificultades para ser medido, pese a los avances en los sistemas de clasificación de pacientes.

En promedio, la atención sanitaria produce más beneficios en cantidad y calidad de vida que costes. Sin embargo, «más» no siempre es «mejor» sino que con frecuencia suele ser «peor».

La sociedad gasta sumas ingentes en el desarrollo de nueva tecnología (900 millones de dólares, por ejemplo, cada medicamento) y solo calderilla en la búsqueda de maneras de utilizar mejor la tecnología de la que ya disponemos. Sin embargo, para conseguir una asistencia de calidad basta aplicar lo que ya se sabe y recordar las constataciones básicas:

Sumario

—Los sistemas sanitarios funcionarán cuando los profesionales sanitarios que asignan los recursos tengan la información y los incentivos convenientes para tomar decisiones clínicas, coste-efectivos.

—La auténtica gestión sanitaria y la auténtica gestión clínica pasan por reducir la brecha entre lo que se podría lograr —con la tecnología y los recursos disponibles— y lo que realmente estamos obteniendo.

—No pagues por A si quieres obtener B. Si pretendes calidad paga por ella. Es cuestión de incentivos entre dos actores: el usuario y el profesional. Algo se puede hacer sobre el primero, mucho sobre el segundo.

En 2001, la publicación del informe del Instituto de Medicina de EEUU confirmó que entre la asistencia sanitaria que tenemos y la que podríamos tener no solo existe un trecho sino un abismo. No falta dinero ni tecnología ni conocimiento. «Más» no siempre significa «mejor»: cómo se gasta (¡la buena Medicina!) resulta más relevante que cuánto se gasta. Pero preocuparse de lo que realmente importa —la efectividad, reducir el abismo en calidad— acarrea incomprensiones, luce poco y consigue escasa financiación. Pese a todo, la motivación intrínseca funciona y en la literatura encontramos publicaciones que informan sobre experiencias de mejora válida, incluso hay publicaciones secundarias —por ejemplo, *Gestión Clínica y Sanitaria*— empeñadas en esta cruzada. Como esquema inspirador sirva el trabajo de Salvador Peiró (2003) acerca de la reducción de la brecha entre eficacia y efectividad en el tratamiento en España del infarto agudo de miocardio.

Del pagar por «ser» al pagar por resultado, pasando por el pagar por «hacer»

En los primeros conciertos del extinto Instituto Nacional de Previsión con hospitales ajenos a la Seguridad Social se partía de la estructura de estos para pactar las tarifas. Desde 1978, el también desaparecido Instituto Nacional de la Salud utilizó como criterio fundamental de pago lo que el hospital «era»: el grado de especialización de sus servicios, la dotación de medios diagnósticos y terapéuticos, la calificación de las plantillas, etc. Se desarrolló incluso una parrilla de categorías —desde básico rural a urbano con mucha superespecialización— en la que cualquier centro hospitalario encajaba. Se pagaba pues por «ser», según la hipótesis implícita de que una mayor complejidad de la oferta iba asociada a una mayor diversidad y gravedad de los enfermos atendidos, hipótesis que obviamente era rechazable en bastantes ocasiones.

Parecía más sensato que los centros sanitarios fueran compensados en función de lo que «hacían», pero, como sabemos, no es lo mismo medir kilowatios o toneladas de acero que estancias o visitas. Fueron los sistemas de ajuste de riesgo los que permitieron clasificar a los pacientes en grupos de isoconsumo de recursos. Si no se sabe medir lo que «hacen» los centros, no se pueden establecer comparaciones ni asignar los recursos (presupuestos, contratos...) a quienes lo hagan mejor. Efectivamente, la carencia de una medida del producto (aunque fuera intermedio) era un problema sanita-

Sumario

rio, como lo era —y continúa siéndolo— en los sistemas de educación, justicia, defensa, policía, o asuntos exteriores. Algunas actividades de las organizaciones que producen directamente para el mercado también se enfrentan a esta dificultad: relaciones públicas, recursos humanos o planificación estratégica; en ocasiones han de recurrir a una modalidad de *benchmarking* conocida como método Bruegel o «ciegos conduciendo a ciegos».

Los Grupos Relacionados con el Diagnóstico (GRD) de Fetter, implantados en EEUU al inicio del año fiscal de 1984 permitieron aproximar mejor la medida de una parte del producto «intermedio» hospitalario (lo que se «hace»), además de ilustrar una rápida transición de la investigación universitaria sobre servicios sanitarios a la aplicación práctica y su rápida difusión. En 1987, el Consejo Interterritorial del Ministerio de Sanidad y Consumo en España aprobó establecer un Conjunto Mínimo Básico de Datos (CMBD) en el alta hospitalaria, de acuerdo con los aceptados tanto por el Comité Nacional de Estadísticas Vitales y de Salud de EEUU, como por la Comisión de la Comunidad Europea sobre el conjunto mínimo básico de datos europeo, lo cual facilitó la extensión de los GRD.

Los Grupos Clínicos Ajustados (GCA) también vinieron de EEUU, esta vez de Johns Hopkins, con el propósito de medir el producto intermedio en atención ambulatoria, y fueron objeto de tempranas validaciones en Andalucía y Cataluña, fundamentalmente. Sin embargo, España continúa sin un CMBD en Atención Primaria.

Tres sistemas de ajuste de riesgo basados en el diagnóstico pueden ser útiles tanto para un pago capitativo de base territorial como para identificar pacientes tributarios de gestión de casos, seguir el estado de salud de las poblaciones y planificar/evaluar los servicios sanitarios: el mencionado GCA, el de los *Clinical Risk Groups* (GRG) de 3M y el de las *Hierarchical Co-existing Conditions* (HCC) de DxCG, citados por el orden de atención que hasta ahora han recibido en España. Antes de optar sensatamente por alguno de los tres, mencionar que este tema se trata, por ejemplo, en un reciente documento de la Fundación Canadiense de Investigación sobre Servicios Sanitarios (Berlinguet *et al*., 2005) que ha orientado a varias provincias hacia la elección de los CGA.

No obstante, «hacer» sigue implicando visitas, exploraciones, intervenciones, etc., algunas de las cuales —como las vinculadas con el internamiento hospitalario— podemos cuantificar aceptablemente y otras —como la consulta ambulatoria de los hospitales, contada con los dedos y sin ajustar— no tanto. En cualquier caso, ha habido una transición paulatina desde premiar por «ser» a compensar por «hacer», y se ha mantenido donde está indicado —cámara hiperbárica, servicio de urgencias rural— el pago por «ser» (pues en este tipo de servicios cuanto menos «hagan» mejor).

Cuando falta todavía mucho para alcanzar medidas de actividad satisfactorias, curiosamente sabemos que ese no es el objetivo. En una sociedad de enfermos crónicos (diabéticos, EPOC, insuficientes cardíacos...), un mayor frenesí asistencial no suele ser indicio de mejor calidad sino todo lo contrario: cuanto más (comas, ingresos no programados...), peor. ¿Por qué financiamos públicamente

Sumario

los servicios sanitarios? Por su impacto en la salud. Pensemos, pues, en pagar por él. De acuerdo, muchas son las variables que influyen en el estado de salud y resulta muy difícil establecer el porcentaje atribuible a la intervención sanitaria. Remontémonos un poco en el tiempo, hasta antes de nacer; si supiéramos que seremos diabéticos o que tendremos un cáncer digestivo ¿en qué país del mundo nos gustaría vivir? En el país que mejor controlara la diabetes —medible a través de la hemoglobina glicosilada— o en el que tuviera el mayor índice de supervivencia tras el cáncer —aproximable, por ejemplo, por el porcentaje de cánceres digestivos diagnosticados en Urgencias Hospitalarias—. De manera que ya se puede decir algo sobre los resultados que realmente importan a los ciudadanos.

El necesario avance hacia la medición de resultados de la atención sanitaria contribuirá a reducir la brecha entre eficacia y efectividad y a obtener una base de evaluación del comportamiento sobre la que incentivar la calidad de los proveedores y la elección informada de los usuarios. Todo con un coste mundial de investigación en servicios sanitarios muy inferior al de desarrollar un nuevo medicamento. «Más» no siempre es «mejor», pero «mejor» (servicio) siempre es «más» (salud).

2. **LA ACTUAL GESTIÓN SANITARIA**

Algunos procesos sanitarios, por su radical impacto sobre la vida, requieren calidad «seis sigmas» (seis desviaciones estándares o un 99,999997% de probabilidad de ausencia de error) y, aún así, no podemos asegurar que la suma de procesos de mucha calidad nos proporcione calidad en el conjunto del sistema sanitario. La excelencia local puede quedar fácilmente sepultada por la mediocridad sistémica.

Para juzgar la calidad de un sistema sanitario es necesario conocer este tipo de datos: la tasa de complicaciones graves en enfermos diabéticos, la supervivencia en cáncer de mama, la prevalencia del tabaquismo, la tasa de resistencias bacterianas, el porcentaje de cánceres digestivos que se diagnostican en urgencias hospitalarias (Porta 1998), la prevalencia de anemias ferropénicas, etc., en general tasas de circunstancias adversas evitables referidas a la utilización, la morbilidad o la mortalidad.

Cerca del 50% de la población española pronto tendrá al menos una enfermedad crónica (asma, depresión, diabetes, insuficiencia cardiaca congestiva...). Existe suficiente evidencia científica para afirmar que la atención efectiva a los enfermos crónicos exige un cambio en las formas de prestar la atención sanitaria (Wagner *et al*. 2001, Ibern 2006).

Desbordar el marco supone pensar —sobre una base poblacional— acerca de la mejor forma de actuar sobre las enfermedades, en todas sus fases, incorporando dos constataciones *novedosas* de los últimos setenta años.

La primera *novedad:* el legado de Walter Shewhart y otros, en la línea del efecto Hawthorne de los años treinta, padres intelectuales de las técnicas de mejora continua de la calidad (Berwick, 1991). El efecto Hawthorne constató que la motivación de las personas importa, que quién sabe más acerca de un trabajo es quien lo está desempeñando y que sería conve-

niente explicar los errores propios para que otras personas no incurran en ellos; por ejemplo, dedicar alguna sesión clínica a pacientes con los que se han cometido errores (Chassin y Becher, 2002). En sanidad, la constatación de variaciones aparentemente arbitrarias en la práctica médica, especialmente en los últimos treinta años, impulsada por Wennberg, ha proporcionado el acicate intelectual y profesional para pensar y actuar en términos de *mejora continua de la calidad* (aunque cada cual lo llame como quiera). Y esto puede realizarse con ayuda de la segunda *novedad*.

La segunda *novedad*: la constatación de Rose (1988) de que las causas de incidencia de enfermedad en una población no coinciden con las que explican la enfermedad en un individuo. Tan importante, o más, que actuar sobre la *letalidad*, o las secuelas mórbidas, es procurar desplazar a la izquierda la curva de *incidencia* de enfermedad. Y en esta disminución de la incidencia de enfermedad, la prevención, fundamentalmente la primaria, desempeña un cometido esencial.

Veamos ahora el estado de ambas *novedades* en España.

Primera novedad: variaciones en la práctica médica en el sistema nacional de salud español

Resulta plenamente compatible el hecho de que la atención sanitaria, en promedio, ocasione más beneficios que costes con el hecho de que existan actuaciones claramente perjudiciales —clínica y económicamente—, así como lagunas de desatención. Tiene poco sentido que hasta un 50% de la atención recomendada deje de prestarse (Glynn *et al*., 2003) cuando el 30% de los gastos de atención sanitaria podrían eliminarse sin afectar negativamente a los resultados en términos de salud (Fisher *et al*., 2003). Dicho de otra forma, el 30% del total de gasto sanitario— 15% del PIB de EEUU y 7,6% del PIB de España— resulta superfluo, y por tanto podría dedicarse a la atención no prestada.

Las variaciones en la práctica médica (VPM) señalan las áreas donde puede haber mayores problemas de calidad y llevan a formular la pregunta clínica y socialmente correcta: ¿para qué (se interviene)?, en lugar de: ¿cuánto (se interviene)? Se trataría de acabar con las políticas sanitarias preocupadas únicamente por el gasto sanitario para adentrarse en el binomio resultados y costes, tanto en el ámbito clínico como en el sanitario.

Atender a la necesidad con financiación pública define la esencia de un servicio nacional de salud. Las VPM nos dicen que esa necesidad **no** guía la utilización, pero sí lo hacen otros factores, como la incertidumbre y la ignorancia, que propician estilos de práctica muy diferentes, influidos de forma muy especial por la *oferta* (más en EEUU que en España) y por la «*fascinación tecnológica*» (ya se verá si se da más en España que en EEUU).

Las actuaciones sobre la *oferta* modelan el sistema sanitario. Ahora se admite la lectura equivocada de la caída del muro de Berlín como un «triunfo del mercado». El libre mercado es la condición necesaria, pero una correcta intervención del Estado es la condición suficiente para que un país prospere. Hay que reivindi-

car la planificación sanitaria de la misma forma que hay que reivindicar una mayor competencia en el sector[1]. Aquí no se trata de saber más. Simplemente de aplicar lo que ya se sabe y planificar, como mínimo igual que otros países «nada intervencionistas».

Más peliaguda se presenta la *fascinación tecnológica*, hipótesis según la cual serían los incentivos de tipo profesional —asociados a la innovación y complejidad de la atención prestada— los que determinarían la utilización sanitaria. Las patologías más habituales, y con tratamientos menos sofisticados, quedarían relegadas en favor de procesos complejos o realizados con tecnologías innovadoras sobre patologías con menor carga de morbilidad (Peiró y Bernal, 2006). La probabilidad de que se resuelva un problema sanitario depende menos de su importancia que de su atractivo (*profile*). En el mundo, el sida es más «atractivo» que la tuberculosis (y más atractivo actuar sobre el sida por medio de anti-retrovirales que a través de procesos mucho más coste-efectivos). En España, la artroplastia de rodilla es mucho más «atractiva» que la reparación del juanete. Poco ha de sorprender que las tecnologías mejor valoradas por los profesionales sanitarios en España sean tecnologías instrumentales como la radioimagen, la angioplastia coronaria o la endoscopia gastrointestinal, en detrimento de las intervenciones más efectivas para reducir la carga de enfermedad (Solá Morales, 2005).

[1] La influencia de la oferta en la utilización sanitaria así como la pérdida de calidad originada por pequeños volúmenes de intervenciones evocan conocimientos sólidamente contrastados.

Las variaciones en la práctica médica (VPM): más allá de la Sanidad

Convenientemente divulgadas, las VPM pueden estimular una mejora constante en el sector sanitario si saben propiciar las respuestas adecuadas (control de oferta, capitación, atención primaria resolutiva, inversión del 4% del presupuesto sanitario en tecnologías de la información y la comunicación (TIC), programas de gestión de enfermedades, etc.), siempre que se tenga conciencia de que un sistema sanitario funcionará bien cuando sus profesionales dispongan de la información necesaria —y la voluntad— para tomar decisiones, fundamentalmente clínicas, coste efectivas. Las VPM pueden incluso «contagiar» a otros sectores (educativo, judicial, seguridad) la actitud intelectual de cuestionamiento constante de la práctica y de escepticismo científico ante los dogmas recibidos, pero deberían servir sobre todo para adoptar una perspectiva más centrada en los pacientes (sin pasarse, pues salud «objetiva» y «subjetiva» no siempre están relacionadas) que formara parte de la política general del país.

Y esta es la batalla más difícil: la mayoría de los que toman decisiones económicas y los creadores de opinión desconocen un sistema nacional de salud que suelen utilizar privilegiadamente, con la ayuda de recomendaciones e influencias. Cuando la mitad de las mejoras en el bienestar de un país desarrollado se derivan de las ganancias en cantidad y calidad de vida, cuando la biotecnología constituye un sector estratégico, y cuando debe compatibilizarse la mejora de la productividad en España con el sostenimiento del estado de bienestar, no cabe que las po-

Sumario

líticas económico-industriales y sanitarias sigan ignorándose (Ortún *et al*., 2005).

Variaciones de la práctica médica, ¿cómo sostener una carrera de fondo?

Al igual que Hsiao (1988) contribuyó a recuperar «la silla» de Marañón como mejor instrumento diagnóstico (profilaxis de la «fascinación tecnológica» a través de cambios en los incentivos), necesitamos evitar que programas tan encomiables como *Ingenio 2010* (el instrumento prioritario en políticas públicas de I+D+i) caigan en una eventual «fascinación tecnológica». Tanto en investigación como en sanidad importa más cómo se gasta que cuánto se gasta. *Cénit*, para aumentar la colaboración público-privada, *Consolíder*, para reforzar masa crítica y excelencia investigadora, y *Avanz@*, nuevos proyectos y servicios en TIC, son los tres programas estratégicos de *Ingenio 2010*. El impulso biotecnológico a través de la colaboración público-privada que contempla el *Cénit* requiere algo que con frecuencia se olvida: una colaboración pública-privada entre economía, industria y sanidad. En pocos sectores están menos desarrolladas las TIC que en sanidad, uno de los sectores que más trabaja con información. Será, no obstante, en el marco de los Consorcios de Investigación Biomédica on Red (*Ciber)* donde tendrá continuidad una investigación como la del grupo de Variaciones en la Práctica Médica de la Red de Investigación en Resultados y Servicios de Salud (VPM/IRYSS),[2] que requiere perseverancia, masa crítica y excelencia. La investigación sobre VPM no precisa grandes instalaciones científico-tecnológicas, pero su impacto en la práctica clínica puede superar claramente a las habituales innovaciones financiadas fundamentalmente por las empresas.

La *sostenibilidad* de la componente sanitaria del estado de bienestar no constituye especial problema: 17% del aumento en recaudación durante los próximos diez años sin variar la presión fiscal (Puig *et al*., 2004). El problema radica en la *deseabilidad*: en que los beneficios en términos de salud superen sus costes, lo cual supone atacar tanto el problema del consumo útil que no se proporciona como el del 30% que puede ser totalmente innecesario.

Se trata, en palabras del coordinador científico del proyecto VPM-IRYS (Bernal, 2005), no solo de servir a la comunidad científica, sino también —y fundamentalmente— a la toma de decisiones de gestión y política sanitaria para incrementar tecnologías efectivas allí donde haya infrautilización de las mismas, reducir la sobreutilización de servicios inefectivos o inseguros, incorporar las preferencias de los pacientes, gestionar la oferta de recursos, y propulsar políticas públicas más eficientes. Se trata de compatibilizar, en suma, la mejora de la productividad en España con el perfeccionamiento de su estado de bienestar; algo que los países del Norte de Europa demuestran en su práctica diaria que es posible.

La segunda novedad

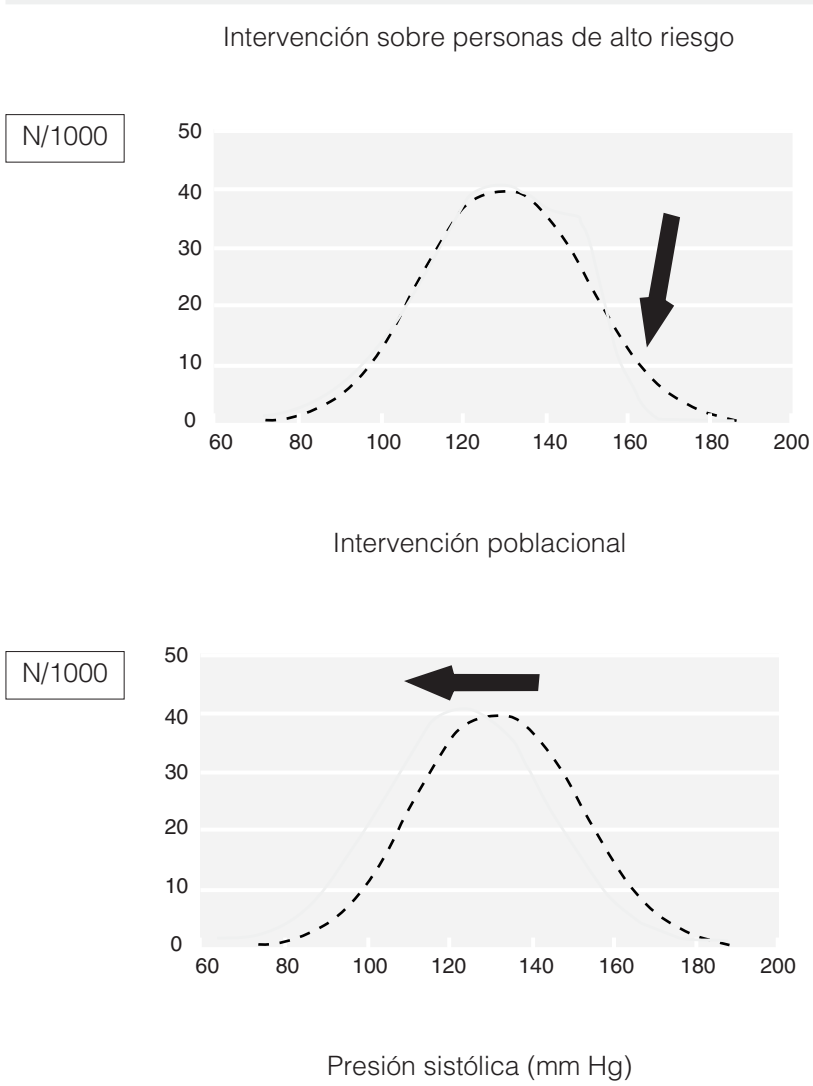
Los beneficios de las intervenciones sanitarias pueden estimarse a partir de la

[2] IRYSS corresponde a la Red de investigación cooperativa para la Investigación en Resultados de Salud y Servicios Sanitarios.

Sumario

Gráfico n.º 1

Intervención sobre personas de alto riesgo comparada con la intervención población: Idealización de los efectos sobre la distribución de la presión arterial sistólica.



disposición a pagar, observada o declarada, de los individuos por una mejor salud. La revisión de decisiones así como diversos estudios relevantes nos llevan a proponer para España una cifra guía de 30.000 euros por año de vida ajustado por calidad (AVAC) como umbral de eficiencia. Precisamente la valoración de los

AVAC permite afirmar que en estas tres últimas décadas cada peseta gastada en intervenciones sanitarias ha resultado beneficiosa en promedio, aunque también es cierto que los gastos alternativos —en cambios de estilos de vida por ejemplo— podrían haber sido más eficientes.

Muchos parámetros fisiológicos tienen distribuciones normales. La tensión arterial, tanto sistólica como diastólica, también. La hipertensión es un importante factor de riesgo de diversas enfermedades. Su control puede realizarse tratando de desplazar a la izquierda la distribución de enfermedad y/o actuando clínicamente sobre la cola de la derecha de la distribución (los hipertensos).

En la gráfico 1 se representa el impacto previsible sobre la distribución de la tensión arterial de ambos tipos de intervenciones. Según datos de EEUU referidos a las cuatro últimas décadas sabemos que los antihipertensivos administrados a la población de riesgo han sido una de las intervenciones más efectivas: se ha ganado hasta medio año de vida en promedio por persona (Long *et al.*, 2006). Ello no implica que la medicación sea siempre la manera más eficiente de disminuir la mortalidad y la morbilidad derivada de las enfermedades del aparato circulatorio (se puede actuar de varias formas y sobre los distintos factores de riesgo). En el seguimiento realizado entre mediados de los ochenta y mediados de los noventa, sobre 38 poblaciones de 21 países desarrollados, se reveló que las mejoras llevadas a cabo en salud pública resultaron más poderosas que la medicación, en la disminución habida en todos los centiles de la distribución de la tensión de la población, (Tunstall-Pedoe *et al.*, 2006).

3. CARACTERÍSTICAS DIFERENCIADORAS DE LAS ORGANIZACIONES SANITARIAS

Podemos distinguir cinco características específicas de las organizaciones sanitarias.

Primera. Tradicionalmente se ha hablado de *hospitalcentrismo* pero puede generalizarse el concepto a *organización-centrismo* y *nivelcentrismo* en numerosos sistemas sanitarios, que favorecen tanto la competencia entre niveles como el hecho de que los centros no consideren los efectos externos —positivos o negativos— que una parte importante de sus decisiones tienen en otros centros sanitarios. El *nivelcentrismo* lleva a optimizar el comportamiento en cada nivel, lo que fácilmente puede suponer:

- No tener en cuenta los beneficios que puedan redundar en otros niveles asistenciales, por ejemplo: ¿por qué prescribir una medicación que reduce las hospitalizaciones pero cuesta más?
- Derivar, compensando previamente cuando haga falta, los casos que no interesan. *Acicalar y largar* en la terminología de Shem (1998).
- Retener los casos que desde cualquier perspectiva, ya sea por interés científico o por comodidad, parezcan más interesantes.
- No preocuparse por la adecuación de las derivaciones procedentes de otro nivel asistencial (o de la adecuación de las pruebas diagnósticas solicitadas por otros servicios dentro del mismo nivel asistencial).

Sumario

—Duplicar las pruebas, con su correspondiente iatrogenia.

No se estimulan, por tanto, los comportamientos cuyas mejoras se reflejan en otros niveles asistenciales.

Segunda. Importancia de las normas clínicas como institución. Las normas clínicas comprenden tanto el conjunto de comportamientos que la profesión estima aceptables (y que sanciona con reconocimiento u ostracismo) como el conjunto de valores y expectativas compartidos por el grupo profesional de referencia. Las normas clínicas constituyen una garantía de que se hará lo mejor para los pacientes con independencia de otras influencias. Eso sí, la fidelidad a la profesión resulta con frecuencia más fuerte que la fidelidad a la organización para la que se trabaja.

Tercera. Coordinación muy intensa dentro y entre organizaciones sanitarias. Al igual que ocurre con el diseño de un avión, el mejor diseño no se obtiene encargando el diseño de sus componentes (fuselaje, navegación, motor, mobiliario…) a los mejores especialistas y reuniendo los diseños después; el avión no volaría. Se requiere una relación constante —prácticamente cotidiana— entre los distintos grupos encargados del diseño de cada parte, lo cual imposibilita la externalización, pues no resulta practicable la renegociación permanente de los contratos. La alternativa que mejor responde a esta necesidad de coordinación muy intensa pasa por algún tipo de organización jerárquica.

Cuarta. Ausencia de soluciones fáciles para los problemas de incentivación. Ni sirve la búsqueda de lucro, sin limitaciones, ya que cortocircuita las decisiones clínicas en interés del paciente, ni puede ignorarse las características de la sociedad en la que vivimos.

Quinta. Aplicación muy atenuada de la regla de oro del control interno: la separación entre gestión de la decisión (preparación de alternativas y ejecución de la elegida) y control de la decisión (elección entre alternativas y evaluación de su ejecución) que ha de producirse siempre que quien toma las decisiones no soporta plenamente las consecuencias de las mismas, no tiene lugar en las decisiones clínicas.

Entre monopolios cabe la competencia por comparación

El monopolio constituye un caso específico de economías de escala ya que pueden producirse los servicios requeridos por un mercado geográficamente establecido a menor coste y mayor calidad cuando hay un único proveedor en lugar de dos o más.

Sin embargo, los monopolios suelen resultar perjudiciales para el bienestar social, de ahí que la política deba evitar su aparición o, como mínimo, su persistencia temporal. La política de una empresa, en cambio, busca conseguir tanto *poder de mercado* como pueda, es decir, tiende a parecerse lo más posible a un monopolio.

El caso es que la delimitación óptima de los servicios sanitarios, como servicios personales, en mercados relevantes se realiza a partir del establecimiento de ámbitos geográficos que cumplen dos condiciones: primera, los proveedores del ámbito-mercado establecido han de prestar como mínimo el 75-90% de los

servicios que reciben los usuarios de ese ámbito-mercado; segunda, el porcentaje de servicios que los residentes de ese ámbito-mercado reciben de proveedores situados en otros ámbitos no ha de exceder el 25-10% (Dalmau y Puig, 1998).

La aplicación de estos criterios en los países desarrollados nos muestra como —excepto en las zonas de alta densidad demográfica— los proveedores sanitarios constituyen un monopolio o bien están radicados en ámbitos-mercados muy concentrados (con pocos proveedores).

De ahí que se hayan propuesto diversas formas atenuadas de competencia. Y aquí conviene recordar un resultado derivado de la comparación de diversas formas organizativas en sanidad y en otros sectores (privadas, con y sin finalidad lucrativa, y públicas): el grado de competencia en el mercado de servicios afecta más al comportamiento de una organización que la naturaleza, pública o privada, lucrativa o no, de su propiedad. La bondad y capacidad innovadora de cualquier organización depende más de los estímulos que reciba —o de los obstáculos que deba superar— que de quiénes ostenten su propiedad (Callejón *et al*., 2002).

Por muy monopolio que sea una organización sanitaria, nunca lo será más que la Agencia Tributaria y sin embargo nada impide a la Agencia comparar la eficiencia recaudatoria de sus inspectores ajustando por sus respectivas bases impositivas. Alguien que se sabe comparado, tomando en consideración, por supuesto, las variables que él no controla, estará más cercano al grado de competencia socialmente saludable que el inimpugnable vitalicio.

Esta competencia por comparación resulta válida tanto para las políticas de salud, estilo OMS, como para la gestión de centros, estilo *los mejores hospitales* (Peiró 2002), o la gestión clínica. Sus efectos benéficos han sido documentados, como mínimo, para esta última (Chassin *et al*., 1996).

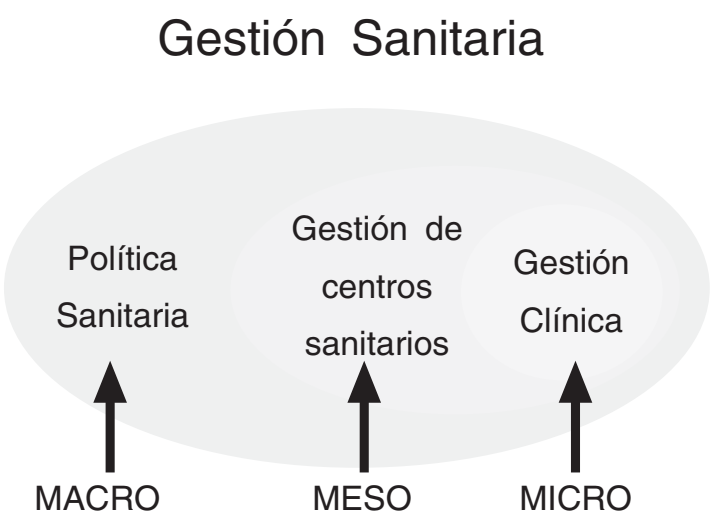
4. LA DIFUSA FRONTERA ENTRE POLÍTICA Y GESTIÓN

Entre los ámbitos de la gestión sanitaria (Figura 2) es en el de la macro-gestión o política sanitaria, en el que resulta más difícil delimitar dónde puede acabar la política y empezar la gestión. En sanidad, la política ha de servir para establecer prioridades y alcanzar compromisos entre objetivos sociales contradictorios y que afectan de manera diferente a los distintos actores sociales. La gestión, que pasa por coordinar y motivar a las personas para conseguir los objetivos de una organización, requiere que estos objetivos sean claros y conocidos. Tan perjudicial resulta suplantar la política (Franco decía que él no se metía en política) como invadir la gestión.

Países como Nueva Zelanda pueden permitirse un director del Servicio Nacional de Salud extranjero elegido tras convocatoria pública internacional. En otros países, más próximos a nuestra tradición, cualquier gestor de una organización sanitaria ha de estar, como mínimo, *bien visto* por la autoridad política de la que provendrá la mayor parte de su financiación. Un gestor necesita *cromos* para intercambiar con sus unidades operativas y de apoyo, si quiere que lo ayuden a lograr una serie de objetivos y actividades

Gráfico n.º 2

Los tres tipos de Gestión Sanitaria



que se expresan en un presupuesto. Sin *cromos* poco juego hay. De ahí la lógica del *nihil obstat, nomenatur*, del cribado previo al nombramiento.

Son varios los condicionantes políticos que se imponen a las reformas de la gestión pública sanitaria:

— *Statu quo* y competencia, por leve que sea, resultan incompatibles. No puede figurar en la misma agenda política el mantenimiento de la situación con la introducción de una competencia siempre schumpeteriana, creativa y destructiva al mismo tiempo. Los datos asistenciales y económicos de los que se dispone en la actualidad permiten ya una competencia entre hospitales de agudos mayor de la que parece políticamente deseable (la que se está dando).

— Mayor flexibilidad sin control conduce fácilmente a un pacto entre gestores públicos y sindicatos a expensas del bienestar común. Hay que respetar el binomio libertad /responsabilidad.

— ¿Restricción por el Derecho Administrativo o restricción política? Intereses corporativos y sindicales no sólo pueden ser un freno a la exter-

nalización sino que también pueden limitar la desburocratización del sector público. El diagnóstico de restricción de la discrecionalidad gestora por el Derecho Administrativo lleva a estrategias de flexibilización de los vínculos funcionariales y estatutarios que pueden resultar absolutamente inocuas si el origen de la falta de flexibilidad no está en el derecho administrativo sino en otros factores de índole política. En la historia de las empresas públicas españolas, laboralizar, por ejemplo, puede relacionarse con mayores salarios, en lugar de con mayor productividad, en la historia de las empresas públicas españolas.

— Los juegos políticos, a diferencia de lo que ocurre en el mundo de los negocios o en el comercio internacional, son frecuentemente juegos de suma cero, de competencia destructiva. Los escaños que uno gana, otro los pierde. Esto dificulta la consecución de pactos entre fuerzas políticas, convenientes para la implantación de mejoras en la gestión pública.

5. **RECAPITULACIÓN**

Las características del proceso sanitario aconsejan incentivos débiles y coordinación muy intensa en sus formas organizativas. Por otra parte, para que un sistema sanitario funcione de manera satisfactoria se precisa que quiénes asignan la mayor parte de los recursos en el sector, fundamentalmente los clínicos con sus decisiones diagnósticas y terapéuticas, tengan el conocimiento y las ganas de decidir de manera coste. El resultado implica situar la gestión clínica en el centro del escenario y procurar una evolución de la medicina cada vez más basada en la efectividad, orientada hacia el usuario y preocupada por el mejor uso alternativo de los recursos de que dispone.

Existe conocimiento y experiencia suficiente para ir introduciendo mejoras en los sistemas sanitarios. Tampoco faltan habilidades analíticas que sepan aportar soluciones a los distintos problemas. Sí hay un déficit, en cambio, de habilidades implantadoras derivado, en gran parte, del desequilibrio entre las limitadas satisfacciones y las elevadas demandas que producen los puestos de gestión sanitaria.

Los aspectos clínicos de la constatación bautizada como «gestión clínica» funcionan de forma satisfactoria. No puede decirse lo mismo, sin embargo, de la plasmación organizativa que la citada constatación debería suponer. No habrá buena gestión sanitaria sin buena medicina, pero tampoco puede existir buena medicina sin buena gestión sanitaria. Tanto la gestión sanitaria como la clínica comparten la misma misión: reducir la brecha entre efectividad y eficacia. Pese a la obviedad de la afirmación, no suele plantearse así ni la actividad clínica ni la actividad gestora. En el texto se ha ilustrado con algunos ejemplos el significado de esa brecha.

Actualmente estamos viviendo la transición que lleva del cobrar por *ser* al cobrar por *hacer*. Aunque este parezca menos insatisfactorio que el primero, no constituye el objetivo final: hay que evolucionar hacia un método de cobrar en función de la efectividad y la calidad, acabando con la actual paradoja de *cuanto*

peor (se controla un enfermo crónico) *mejor* (mayores ingresos hospitalarios se producen).

Para proporcionar una visión de conjunto, antes de plantearse quién hace qué en una organización determinada, se conjugan dos *novedades* en los últimos setenta años: la constatación de variaciones aparentemente arbitrarias en la práctica, la entronización del sentido común en los programas, y actitudes, de mejora continua de la calidad, y la distinción entre causa de incidencia y causa de enfermedad.

La buena gestión clínica, para cualquiera de sus profesiones, pasa por una excelente base disciplinar y un largo aprendizaje. La gestión clínica ha de desbordar, no obstante, los límites de los centros, ya que: en primer lugar, una parte importante de las decisiones que en ellos se adoptan producen efectos —perjudiciales o beneficiosos— en otros centros; en segundo lugar, y recíprocamente, la efectividad en la actuación de cualquier organización viene condicionada por el comportamiento del resto de organizaciones sanitarias a las que también acuden sus pacientes.

REFERENCIAS BIBLIOGRÁFICAS

BERLINGUET M, PREYRA C, DEAN S (2005): *Comparing the value of three main diagnostic-based risk-adjustment systems*. Canadian Health Services Research Foundation. Accesible en: http://www.chsrf.ca/final_research/ogc/pdf/berlinguet_final.pdf

BERNAL E (2005): *El atlas VPM*. Atlas Var Prac Med Sist Nac Salud; 1(1): 3-4.

BERWICK D (1991): *Controlling variation in health care: A consultation from Walter Shewhart.* Medical Care; 29 (12): 1212-25.

CALLEJÓN M. ORTÚN V. (2002): «Empresa privada en Sanidad: Grado de adecuación a las diferentes funciones sanitarias». En G. López (ed): *El interfaz público-privado en Sanidad.* Barcelona: Masson, p 103-121.

CHASSIN M, HANNAN E, DEBUONO B (1996): *Benefits and hazards of reporting outcomes publicly.* New Engl J Med; 334: 394-398.

CHASSIN M, BECHER E (2002): *The wrong patient.* Ann Intern Med 2002; 136: 826-833.

DALMAU-MATARRODONA E, PUIG-JUNOY J (1998): *Market structure and hospital efficiency: evaluating potential effects of deregulation in a National Health Service.* Review of Industrial Organization; 13: 447-466.

GLYNN. E. ET AL. (2003): *The quality of health care delivered to adults in the United Status.* New Engl J Med; 348(26): 2635-45.

FISHER E ET AL. (2003): *The implications of regional variations in Medicare spending.* Annals of Internal Medicine; 138(4): 273-287 y 288-298.

HSIAO W ET AL. (1988): *Estimating physicians' work for a resource-based relative-value scale.* New Engl J Med; 319(13): 835-41.

IBERN P, DIR (2006): *Integración asistencial: fundamentos, experiencias y vías de avance.* Barcelona: Masson-Elsevier.

INSTITUTE OF MEDICINE (2001): *Crossing the Quality Chasm: A New Health System for the 21st century.* Washington: National Academy Press.

Sumario

KHATRI GR, FRIEDEN T (2002): *Controlling tuberculosis in India*. New Engl J Med; 347: 1420-5.

LONG G, CUTLER D, BERNDT E ET AL. (2006): *The impact of antihypertensive drugs on the number and risk of death, stroke and myocardial infarction in the United States.* National Bureau of Economic Research Working Paper Series, nº 12096.

ORTÚN V, DIR (2003): *Gestión clínica y sanitaria. De la práctica diaria a la academia, ida y vuelta.* Barcelona: Masson-Elsevier.

ORTÚN V, PUIG-JUNOY J, CALLEJÓN M (2005): «Innovación en medicamentos, precios y salud». En: González López-Valcárcel B (dir): *Difusión de nuevas tecnologías sanitarias y políticas públicas.* Barcelona: Masson-Elsevier, p. 173-193.

PEIRÓ S (2002): *¿Son siempre odiosas las comparaciones? Gestión Clínica y Sanitaria;* 4 (2): 35-37. Accesible en www.iiss.es/gcs/index.htm

PEIRÓ S (2003): *De la gestión de lo complementario a la gestión integral de la atención a la salud*. En Ortún V (dir): Gestión Clínica y Sanitaria. Barcelona: Masson-Elsevier; p 17-87.

PEIRÓ S (2006): «¿A qué incentivos responde la utilización hospitalaria del Sistema Nacional de Salud?» En: Palomo L *et al*. (eds): Informe SESPAS 2006. *Los desajustes en la salud en el mundo desarrollado*. Gaceta Sanitaria 20 (Suplemento 1): 110-116.

PORTA M, FERNÁNDEZ E, BELLOC J, MALALTS N, GALLÉN M, ALONSO J (1988): *Emergency admissions for cancer: a matter of survival?* British Journal of Cancer; 77: 477-484.

PUIG-JUNOY J, LÓPEZ-CASASNOVAS G, ORTÚN V, EDS (2004): *¿Más recursos para la salud?* Barcelona: Masson-Elsevier.

RODRIGO T, CAYLÀ J, GALDÓS-TANGÜIS H, JANSÀ J, BRUGAL T, GARCÍA DE OLALLA P (1999): *Evaluación de los programas de control de la tuberculosis de las Comunidades Autónomas en España*. Med Clin; 113: 604-607.

ROSE G (1988): «Individuos enfermos y poblaciones enfermas». En: *El desafío de la Epidemología. Problemas y lecturas seleccionadas.* Washington: Organización Panamericana de Salud (publicación científica 505), p 900-909.

SHEM M. (1998): *La Casa de Dios*. Barcelona: Anagrama.

SOLÁ-MORALES O (2005): *Preferencias y tecnología: La perspectiva de los profesionales.* Economía y Salud. Boletín Informativo; 18(53): 6-7. Disponible en http://www.aes.es/Publicaciones/ES53.pdf

SPETZ J, SMITH MW, ENNIS SF (2001): *Physician incentives and the timing of cesarean sections: evidence from California*. Med Care; 39: 536-550.

TUNSTALL-PEDOE H, CONNAGHAN J, WOODWARD M, TOLONEN H Y KUULASMAA (2006): *Pattern of declining blood pressure across replicate population surveys of the WHO MONICA project,* mid-1980s to mid-1990s, and the role of medication. BMJ; 332: 629-635.

WAGNER E, AUSTIN B, DAVIS C, Hindmarsh M, Shaefer J, Bonomi A (2001): *Improving chronic illness care: translating evidence into action*. Health Affairs; 20 (6): 64-78.

Sumario

Programación y evaluación pública: un triángulo complejo

El desarrollo teórico de la evaluación de políticas públicas se ha centrado en la elaboración de herramientas y metodologías de análisis de programas y proyectos aislados. Los programas públicos actuales están caracterizados por lo que podría denominarse «un triángulo complejo de relaciones» donde se encuentran la múltiple dimensión sectorial que pretende cubrir las intervenciones públicas de esta naturaleza, la heterogeneidad y multiplicidad de los implicados en un programa y las diferentes escalas administrativas encargadas de la gestión y ejecución de las actuaciones y su estructura política. Además, la interrelación entre estos tres elementos añade más complejidad al objeto de la evaluación. Con la intención de mostrar la naturaleza compleja del enfoque de evaluación, en este artículo se han seleccionado dos casos de estudio: la evaluación de la Iniciativa Comunitaria Equal en Andalucía y la evaluación de la política municipal de vivienda en Sevilla.

Politika publikoen ebaluazioaren garapen teorikoak programa eta proiektu isolatuak aztertzeko bitarteko eta metodologietan jarri du arreta. Egungo programa publikoek «harremanen triangelu konplexua» delakoa dute ezaugarri; bertan daude honelako esku-hartze publikoak babestu nahi dituen askotariko dimentsio sektoriala, programa batean inplikatutako alderdien heterogeneotasun eta anizkoiztasuna eta jardunak kudeatu eta betearazteaz eta egitura politikoaz arduratzen diren administrazioko eskalak. Gainera, hiru osagai horien arteko erlazioak are konplexuago bihurtzen du ebaluazio-gaia. Artikulu honetan bi aztergai aukeratu dira, honelako ebaluazio-ikuspegiaren konplexutasuna azaltzeko asmoz: Europako Erkidegoaren Equal ekimenaren ebaluazioa Andaluzian eta udalaren etxebizitza-politikaren ebaluazioa Sevillan.

The theoretical development of the public politics evaluation has been centred on the elaboration of tools and methodologies of analysis of isolated programmes and projects. The present public programmes are characterised by what could be called «a complex triangle of relations» where are encountered the multiple dimension relating to a sector that pretend to cover the public supervision of this nature, the heterogeneous nature and the multiplicity of the concerned in a programme and the different administrative scales in charge of the management and the fulfilment of the actions and its politic structure. Moreover, the interrelation between these three elements adds more complexity to the object of the evaluation. With the aim to demonstrate the complex nature of the evaluation, it has been selected in this article two cases of study: the evaluation of the Equal Initiative Community in Andalusia and the evaluation of the municipal politics of housing in Seville.

José Luis Osuna Llaneza
Carmen Vélez Méndez
Ana Cirera León
Juan Murciano Rosado
Universidad de Sevilla
Instituto de Desarrollo Regional

ÍNDICE

1. Distancia entre los retos teóricos y prácticos de la evaluación de políticas públicas
2. Una relación triangular compleja
3. La interacción sectorial de las políticas públicas
4. La multiplicidad y heterogeneidad de los agentes clave de una política
5. Los niveles administrativos de las políticas públicas
6. Los sistemas de evaluación: una realidad a mejorar
7. De la destilación en cascada a la generación de un sistema de evaluación
Referencias bibliográficas

Palabras clave: evaluación, seguimiento, participación, política, vivienda, iniciativa comunitaria Equal

N.º de clasificación JEL: O2

1. DISTANCIA ENTRE LOS RETOS TEÓRICOS Y PRÁCTICOS DE LA EVALUACIÓN DE POLÍTICAS PÚBLICAS

Los desarrollos teóricos y empíricos en materia de evaluación de políticas públicas se han centrado fundamentalmente en aportar herramientas y soluciones metodológicas a la evaluación de proyectos concretos, considerados de forma individual y aislada. Así, la literatura existente tiene pocas referencias sobre la evaluación de políticas y programas más complejos caracterizados por ser de naturaleza intergubernamental, provocar impactos en distintas esferas del ámbito económico-social y por la diversidad y heterogeneidad de actores involucrados.

Sin embargo, la realidad del proceso político y de toma de decisiones presenta una configuración bien distinta a la de planes aislados e individuales; los programas que se diseñan y ejecutan encajan en la agenda política como una pieza de un puzzle que persigue, o al menos así debe serlo *a priori*, objetivos comunes y más amplios que están intrínsecamente relacionados entre sí.

Si bien este problema de correlaciones entre programas se ha venido cubriendo, al menos en los países donde se han realizado evaluaciones como respuesta a las

obligaciones reglamentarias de los proyectos financiados por la Comisión Europea[1], cuando se transfiere el análisis del ámbito de los fondos estructurales a otros campos la cuestión se hace más compleja.

Y esto es así puesto que el análisis de sinergias que se ha venido exigiendo en las evaluaciones de los programas financiados con fondos europeos no responde a las necesidades evaluativas de otros programas complejos actuales. Estos estudios de sinergias se han basado, casi exclusivamente, en valorar la coherencia entre los objetivos propuestos por diferentes programas bajo un esquema metodológico un tanto encorsetado y de escasa utilidad para la toma de decisiones, puesto que en ocasiones esos análisis han consistido en la elaboración de un *checklist* de complementariedad más que en la búsqueda de las causas del funcionamiento de los procesos implícitos en los programas.

Cuando lo que se evalúa es una política nacional o regional en sentido amplio, donde intervienen distintas unidades administrativas con distintas competencias y con capacidad de decisión diferentes, la evaluación de las interrelaciones entre los diferentes programas elaborados por el gobierno se convierte en una tarea insoslayable. Aún más cuando el beneficiario de la política concibe al programa como un elemento más de lo que su gobierno debe ejecutar, y no como una unidad de programación independiente. De hecho, en ocasiones, el beneficiario último de las actuaciones no conoce el marco programático en el que se circunscribe la actuación.

Así lo han puesto de manifiesto los beneficiarios en diversas encuestas que hemos realizado en el Instituto de Desarrollo Regional (IDR) para valorar el impacto de las actuaciones públicas, en materia de comercio, vivienda, autoempleo y aeronáutica, entre otras. En estas encuestas se corrobora el desconocimiento de la imbricación de las ayudas en su marco de planificación superior (Plan Integral de Comercio Interior de Andalucía; Plan Andaluz de Suelo y Vivienda, etc.) y sobre el órgano que denegaba u otorgaba la ayuda. Para ellos, frente al entramado burocrático disperso, el responsable último es único, y normalmente está centralizado en el gobierno regional.

Todo ello viene lógicamente explicado porque en una estrategia de desarrollo regional la planificación debe apoyarse en diferentes medidas bajo el paraguas de un programa de gobierno integral que además, en el caso andaluz, forma parte de un programa tripartito de negociación y consenso entre los dos sindicatos más representativos, la patronal y el gobierno.

Por otra parte, tampoco se hacen esfuerzos tendentes a clarificar y difundir a la población beneficiaria el contenido y contexto de los programas en los que se insertan las ayudas que se conceden y de los fines perseguidos. Cuando se anuncian los programas se hace en el marco de un discurso político en el que no se especifica claramente ni la estrategia de actuación ni las sinergias con otros planes.

Además, los intentos por reunir en un solo programa todos los frentes que en una misma materia se llevan a cabo en

[1] Lo que los sitúa en la primera fase de construcción de la capacidad evaluativa según la clasificación que resultó de la investigación realizada por el MEANS.

un territorio han añadido complejidad no sólo al diseño y gestión de las intervenciones sino a sus evaluaciones. Por ejemplo, este es el caso de la Iniciativa Comunitaria Equal, que ha pretendido aglutinar en un partenariado estratégico a todos los agentes con competencias e intereses en la promoción de empleo a escala local.

Por todo ello, encontramos que el marco metodológico actual no es lo suficientemente flexible y amplio como para dar respuestas satisfactorias a cuestiones relacionadas con la evaluación de programas donde las redes de interacción y estructuras deben ser necesariamente analizadas, valorando su contribución a los procesos de gobernanza y a la potenciación de las organizaciones y territorios y de las políticas que se llevan a cabo.

Con la intención de mostrar la naturaleza compleja del enfoque de evaluación a este tipo de realidad, en este artículo se han seleccionado dos casos de estudio: la evaluación de la Iniciativa Comunitaria Equal en Andalucía (2001-2002) y la evaluación de la política municipal de vivienda en Sevilla.

2. UNA RELACIÓN TRIANGULAR COMPLEJA

El cambio en el modo de producción que dentro del sistema capitalista se está dando desde la última década del siglo XX no sólo tiene como manifestación su carácter planetario, el ser económicamente muy dinámico y producir un amplio grado de exclusión social, sino que también exige un nuevo marco institucional dirigido por la necesidad de una mayor dimensión cualitativa y cuantitativa de los mercados. Como consecuencia de ello, los procesos de descentralización en la toma de decisiones se generalizan dentro de los estados más modernos.

Lo que descriptivamente tiene una clara lógica interna para el sistema, desde el punto de vista de la aplicación de las políticas y programas y/o actuaciones públicas, provoca un conflicto en términos de descentralización frente a coherencia que coloca a la evaluación como elemento para justificar la primera, en aras a una mayor eficiencia, sin que se desvirtúe el sentido de la segunda.

El caso de la Unión Europea y sus relaciones con los Estados miembros, y el de éstos con sus respectivas regiones y/o entidades locales, son paradigmáticos. En efecto, sin la política monetaria, como consecuencia de la creación de una unión económica y monetaria y el mayor grado de integración supranacional alcanzado en toda la historia de la humanidad, y con la política fiscal condicionada, todo el andamiaje institucional del viejo Estado se resiente.

Así, los gobiernos nacionales, regionales y locales tienden a hacer suyas las políticas, programas y actuaciones públicas sectoriales que, además, la percepción ciudadana así lo manifiesta. En definitiva, quieren diseñar de manera autónoma las actuaciones con fondos de terceros y esa contradicción se vuelve insalvable cuando desde lo sectorial (sectores productivos, cooperación exterior, sanidad, educación, etc.) se pretende dar satisfacción a las necesidades colectivas, la mejora en la distribución de la renta y el pleno empleo que, obviamente, tienen un carácter finalista del que carecen las anteriores.

Sumario

La única salida está, pues, en la demostración de la eficiencia y eficacia en la gestión, análisis que debe hacer la evaluación, pero que inevitablemente tendrá que hacerlo asumiendo tal complejidad. Nuevos tiempos exigen nuevas respuestas desde la planificación económica y también de sus herramientas.

Los programas públicos actuales están caracterizados por lo que podría denominarse «un triángulo complejo de relaciones» donde se encuentran:

a) La múltiple dimensión sectorial que pretende cubrir las intervenciones públicas de esta naturaleza.

b) La heterogeneidad y multiplicidad de los implicados en un programa.

c) Las diferentes escalas administrativas encargadas de la gestión y ejecución de las actuaciones y su estructura política.

Además, la interrelación entre estos tres elementos añade complejidad al objeto de la evaluación. Así, la naturaleza de los impactos provocados por un plan se inserta en una compleja realidad socioeconómica, susceptible de no ser correctamente captada si no se dispone de un marco conceptual preciso que ayude a situar el instrumento frente a la problemática socioeconómica a la que responde.

La evaluación en estos supuestos requiere la aplicación de un cuerpo metodológico que incorpore técnicas capaces de medir, cuantitativa y cualitativamente el alcance de la política diseñada. El diseño de la evaluación debe, necesariamente, incorporar la valoración de aspectos relacionados con la incidencia, complementariedad y sinergias de los planes con otras políticas y programas.

3. LA INTERACCIÓN SECTORIAL DE LAS POLÍTICAS PÚBLICAS

Cuando la Administración desarrolla políticas públicas o presta servicios públicos debe hacerlo por mandato constitucional (en virtud del principio de legalidad), de acuerdo con normas reglamentarias o con rango de ley.

Estas políticas no se desarrollan de forma aislada, sino que son interdependientes y el logro de sus objetivos está condicionado por los efectos sinérgicos y complementarios que se logren con la ejecución de cada una.

El caso de la política de vivienda muestra un claro ejemplo de la diversidad de ámbitos sectoriales que pueden estar afectados. En el caso de España, el acceso a una vivienda digna es además un derecho reconocido en la Constitución Española (art. 47 de la Constitución Española[2]).

Además de los objetivos específicos que los diferentes planes de vivienda han perseguido, las actuaciones de las administraciones se han guiado por cuatro objetivos generales:

— Objetivo económico: mantener el nivel de actividad económica y empleo en el sector de la construcción, que es un sector clave para la eco-

[2] «Todos los españoles tienen derecho a disfrutar de una vivienda digna y adecuada. Los poderes públicos promoverán las condiciones necesarias y establecerán las normas pertinentes para hacer efectivo este derecho, regulando la utilización del suelo de acuerdo con el interés general para impedir la especulación. La comunidad participará en las plusvalías que genere la acción urbanística de los entes públicos.»

nomía nacional (10% del VAB nacional y 11% del empleo total).

— Objetivo social: apoyar preferentemente a los colectivos con salarios medios y bajos, inferiores a 3.5 veces el Salario Mínimo Interprofesional[3].

— Objetivo territorial: lograr la adecuación territorial y social entre la oferta y demanda de vivienda actuando en las áreas prioritarias con fuertes intervenciones en el sector público y protegido.

— Objetivo instrumental: buscar el consenso y coordinación entre las administraciones involucradas, especialmente con los ayuntamientos.

En el caso de la Iniciativa Equal en Andalucía, dirigida a unos colectivos que se encuentran fuera de los círculos sociales tradicionales y que tienen escaso o nulo acceso a los programas y servicios públicos de apoyo a la inserción y reinserción laboral, las medidas de política de empleo e inserción abarcan realidades complejas que precisan de la consideración de otras políticas. Entre ellas la social y las transversales de igualdad de oportunidades, de medio ambiente y de nuevas tecnologías de la información y la comunicación, en línea con las prioridades horizontales del Fondo Social Europeo:

— Igualdad de oportunidades: constituye un principio que debe inspirar los proyectos Equal, contemplándola de forma transversal desde la planificación y gestión hasta su seguimiento y evaluación.

— Medio ambiente: los proyectos disponen de un abanico de orientaciones desde Equal que les encaminan hacia la consideración del impacto ambiental de sus intervenciones, sensibilizar a favor de su protección y reforzar el potencial de generación de empleos de futuro relacionados con este sector.

— Nuevas tecnologías de la información y la comunicación (NTIC): para abordar una estrategia de desarrollo de la sociedad de la información desde Equal se recomienda concentrar los esfuerzos en la inclusión de las NTIC en la gestión interna de los proyectos así como en potenciarlas como un nuevo yacimiento de empleo.

Asimismo, el eje de actuación que se concentra en la adaptabilidad de trabajadores y empresas ha provocado en Andalucía el desarrollo de proyectos enmarcados en las políticas sectoriales de ámbito regional como son la promoción del comercio minorista y la adaptación hacia nuevas formas de producción agrícola. Por otro lado, el proyecto andaluz encuadrado en el campo de lucha contra el racismo y la xenofobia ha provocado la inclusión de medidas de política migratoria.

Por tanto, las políticas deben entenderse en el marco de unos objetivos más amplios de la administración. No deben tratarse las políticas de manera aislada, porque están presentes y afectadas por varios ámbitos de actuación pública y en ese sentido deben contribuir a otros fines como el crecimiento económico y del empleo (política de pleno empleo), a la políti-

[3] Actualmente es el Indicador Público de Renta de Efectos Múltiples (IPREM) el que sirve de referencia del nivel de renta para acceder a determinadas prestaciones, beneficios o servicios públicos o para determinar la cuantía de las ayudas.

ca de cohesión social y protección de los colectivos más desfavorecidos (como jóvenes, inmigrantes, familias de ingresos medios o bajos, la tercera edad y discapacitados), apoyar la política de protección a la familia, etc. Todos ellos deben ser considerados en la evaluación.

4. LA MULTIPLICIDAD Y HETEROGENEIDAD DE LOS AGENTES CLAVE DE UNA POLÍTICA

En relación con la heterogeneidad de los agentes implicados, la puesta en marcha de un plan implica la participación de un conjunto diverso de instituciones, gestores y beneficiarios. Así, la eficacia de las actuaciones diseñadas depende, en buena medida, de la adecuación de los mecanismos de gestión y coordinación a la demanda social existente, es decir, a los intereses de estos agentes e instituciones y a las necesidades reales de la población objetivo a la que va destinada la intervención.

La evaluación, en este sentido, debe propiciar una visión de conjunto sobre la coherencia de dichos intereses con los objetivos perseguidos por el plan. El papel del evaluador en este marco debe estar orientado a valorar la coherencia del proceso de transformación del discurso de intenciones en herramientas eficaces y eficientes para la ejecución exitosa de la actuación pública. El evaluador actúa aquí como un facilitador en el diseño de las políticas.

La aplicación de una metodología de evaluación participativa, basada en la configuración de una estructura externa de apoyo al proceso evaluador a través de grupos de trabajo, nos ha llevado a dividir a los agentes implicados en la política de vivienda en tres grandes grupos: beneficiarios, ejecutores y gestores e instituciones.

En el gráfico n° 1 se representan estos agentes que actúan en el diseño, ejecución, gestión y evaluación de la política de vivienda. Esta multiplicidad de organismos exige una planificación eficaz y eficiente que logre conciliar intereses y coordinar sus actuaciones. En la misma línea, la evaluación debe concebirse desde un esquema participativo, aunando las voces de los distintos grupos, colocándose como interlocutor y suministrador de información, no sólo a los poderes públicos y a los gestores de la política, sino también a la sociedad en general.

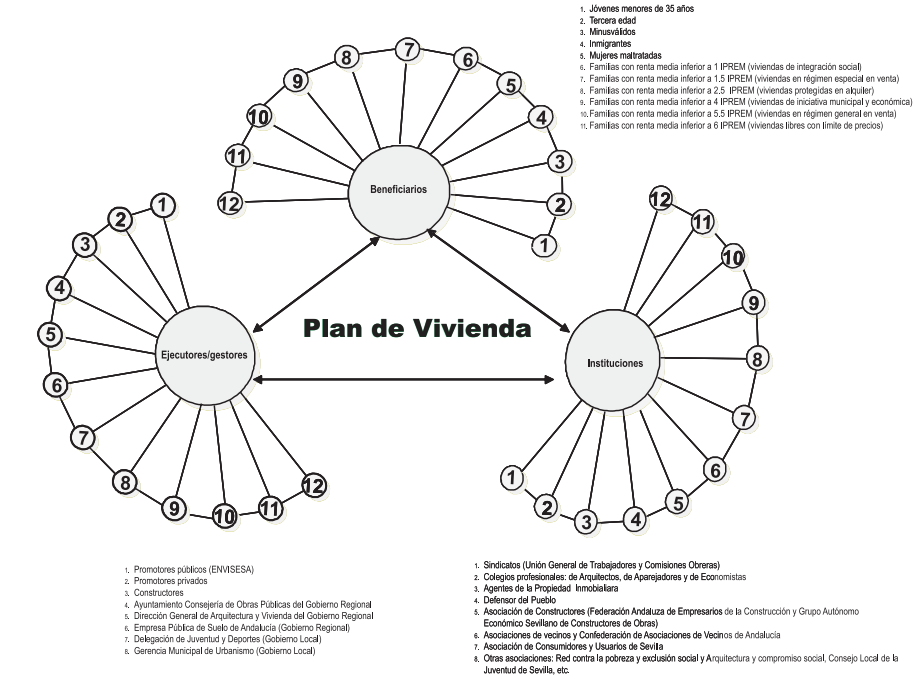
En la IC Equal, mientras que el dibujo del abanico de beneficiarios no presenta mayores problemas (discapacitados, mujeres, inmigrantes, ex-reclusos, ex-toxicómanos, parados de larga duración, etc.), a diferencia de lo que ocurre en el caso del Plan de Vivienda, en el intento de graficar los otros dos, instituciones y ejecutores/gestores, las interrelaciones que los principios rectores y orientaciones establecidos por la Comisión para la Iniciativa Comunitaria establece, añade otra variable a considerar.

En los proyectos Equal la Comisión impone un modelo basado en la cooperación institucional y en la creación de redes locales y transnacionales, de carácter sectorial o geográfico en función del análisis del origen de las discriminaciones en el mercado de trabajo y del ámbito de actuación al que dirigen sus intervenciones.

Así, se han formado Agrupaciones de Desarrollo en las que participan instituciones públicas y entidades privadas, inter-

Grafico n.º 1

Multiagentes. El caso del Plan de Vivienda en Sevilla



Fuente: Elaboración propia

locutores sociales, organizaciones no gubernamentales, empresas y entidades que aportan sus recursos y experiencia de trabajo con los colectivos más vulnerables, cooperando para experimentar nuevas soluciones a los problemas de discriminación y desigualdad en el mercado laboral.

Además de participar en la defensa de sus intereses institucionales, todos estos agentes tienen funciones relacionadas con el diagnóstico, diseño, gestión, ejecución y, en su caso, evaluación, de los proyectos en los que se plasma Equal. Es esta particularidad lo que dificulta deslindar a gestores y ejecutores del resto de instituciones con intereses en los proyectos.

En total son 256 entidades las que participan en las Agrupaciones de Desarrollo que gestionan los 21 proyectos Equal que representa el Gobierno Regional Andaluz. Algunas de ellas participan en más de un proyecto. En esta red participan además 42 agrupaciones europeas, fundamentalmente de Italia, Francia y otros países no mediterráneos.

Sumario

Cuadro n.º 1

Entidades participantes en las Agrupaciones de Desarrollo

Tipo de entidad		Nº de veces que participa
Asociaciones	De empresarios	51
	De trabajadores	21
	De mujeres	28
	Representantes de discapacitados	19
	Otras (fundaciones, asociaciones de barrio, ONGs, y representantes de colectivos marginados)	57
Administraciones Públicas	Gobierno regional	32
	Gobierno provincial	11
	Agrupaciones de municipios	9
	Gobierno local	84
Universidades		8
Institutos de investigación		5
Total		325

- Las asociaciones representan el 55% de los socios de la IC Equal en Andalucía, y de ellas las que tienen mayor representación han sido las de mujeres (en 13 proyectos), discapacitados (en 9 proyectos) e inmigrantes (en 8 proyectos), siendo los sindicatos los que han participado en más proyectos.
- La implicación de los ayuntamientos en las agrupaciones pone de manifiesto la vinculación de los proyectos a los entornos locales.

Fuente: Sistema Equal GSE. Servicio Andaluz de Empleo. Junta de Andalucía

La conformación del número de miembros de las agrupaciones no ha seguido un criterio preestablecido, lo que ha provocado en ocasiones grandes divergencias en el tamaño de la red, más de 40 socios frente a sólo 4 en otros proyectos, situándose el tamaño medio en 14.

Todo este proceso de implicación pretende, en definitiva, como Hemmati apunta (Hemmati, E. 2002):

— Incorporar los intereses y opiniones de una representación equitativa de todos los implicados en un programa o política, tanto para el proceso de toma de decisiones como para reforzar la transparencia en la comunicación y rendición de cuentas entre ellos.

— Fortalecer los principios de transparencia y participación en el proceso democrático.

— Desarrollar partenariados y fortalecer las redes entre los implicados.

Huelga decir que la complejidad y la relevancia de estas intervenciones requieren un sistema que proporcione información sobre todas las fases y actores del plan que se ajuste lo máximo posible a la realidad que se pretende transformar. La evaluación debe permitir satisfacer las exigencias de información de la sociedad, de los órganos representativos del poder ejecutivo y, especialmente, del centro gestor del gasto, dotándolo de un instrumento eficaz para la administración del plan. Información y transparencia son, por tanto, los principios que deben guiar los trabajos de análisis que en el marco de la evaluación de programas se realicen por parte de los evaluadores. Pero no sólo debe ser transparente, para que sea eficaz, el sistema debe proporcionar información en tiempo y forma oportunos, ofreciendo cuadros de pilotaje que permitan tomar decisiones casi en tiempo real.

Es en el diseño de este sistema en el que los evaluadores deben jugar un papel importante en lo que podríamos denominar «determinación de un sistema de seguimiento interno asistido». Además, es fundamental que se haga previo al inicio del plan para que pueda ser llevado a la práctica y que sea lo suficientemente ágil para que se inserte en la rutina diaria de quien recopila y procesa la información que la puesta en marcha de las actuaciones conlleva.

No debemos olvidar que de la evaluación se debe extraer un efecto experiencia sobre buenas prácticas, sobre todo si los planes son un continuo dentro de una línea estratégica de actuación. Por ejemplo, en el caso de la política de vivienda, iniciada en 1991, actualmente se encuentra en ejecución el cuarto período programático. El análisis se utiliza en este marco para generar conocimiento que favorezca el proceso de retroalimentación para sucesivas actuaciones públicas en este ámbito y mejorar las actuaciones que están en vigor, constituyéndose como un *input* elemental para el aprendizaje organizacional en materia de gestión y ejecución de políticas públicas.

5. LOS NIVELES ADMINISTRATIVOS DE LAS POLÍTICAS PÚBLICAS

Además de la multiplicidad de efectos sectoriales sobre los que incide un plan y que le afectan de manera directa e indirecta, y de los distintos agentes que están implicados en alguna de las fases del ciclo de vida de un programa, la evaluación debe considerar la articulación de las distintas instituciones públicas que confluyen. Esto es, la estructura multinivel de la organización competencial de la administración pública.

En el caso de la política de vivienda, la articulación es la siguiente:

— Al Estado, en el ámbito de las competencias relativas a las bases de planificación económica, le corres-

ponde la formulación de planes plurianuales de vivienda. En estos planes, la Administración central determina las normas básicas de acceso y financiación de la vivienda pública y coordina la colaboración de los restantes agentes del sector.

— A las comunidades autónomas, en ejercicio de sus competencias, les corresponde regular en materia de urbanismo y política de suelo, así como aprobar los planes de vivienda de sus respectivos territorios[4].

— Las corporaciones locales son competentes en la formulación y ejecución del planeamiento urbanístico y residencial como titulares del patrimonio municipal de suelo.

Algunos ayuntamientos, como el de Sevilla, se han implicado directamente en los procesos urbanísticos con la creación y oferta de suelo, la cesión de éste para actuaciones públicas de vivienda, la promoción y rehabilitación de viviendas, y la gestión del parque público.

En Equal el esquema general lo diseña, en principio, la Comisión Europea y, tras la discusión y negociación con el conjunto de los Estados miembros y el Parlamento Europeo, adopta la forma de comunicación por la que se aprueban las orientaciones que deben tenerse en cuenta para la elaboración de los correspondientes programas nacionales.

Tanto el Programa de Iniciativa Comunitaria (PIC) como los proyectos que se financien dentro de él han de ser complementarios al conjunto de los recursos y políticas nacionales, regionales o locales en materia de políticas de empleo e inserción socio-laboral de personas que sufran discriminación o desigualdad en el mercado de trabajo. La articulación responde al siguiente esquema:

— La Comisión de la Unión Europea mediante el Fondo Social Europeo instrumenta la IC Equal para el desarrollo de la estrategia comunitaria de inclusión e igualdad recogida en el Tratado de Ámsterdam y en la Estrategia Europea para el Empleo (EEE), que se apoya en cuatro pilares: capacidad de inserción profesional, espíritu de empresa, adaptabilidad e igualdad de oportunidades.

— El Estado, sobre la base de las orientaciones establecidas por la Comisión y ratificadas por el Parlamento Europeo, negocia un Programa de Iniciativa Comunitaria coherente con el Plan Nacional de Acción por el Empleo que cada gobierno debe diseñar[5].

— El Gobierno Regional mediante el Servicio Andaluz de Empleo (SAE), adscrito a la Consejería de Empleo y Desarrollo Tecnológico, participa en el Programa de Iniciativa Comunitaria[6]. Este organismo es competente

[4] La Comunidad Autónoma de Andalucía tiene recogida expresamente esta competencia en el artículo 13.8 del Estatuto de Autonomía.

[5] En España la gestión de las ayudas del FSE está encomendada a la Unidad Administradora del Fondo Social Europeo (UAFSE), adscrita al Ministerio de Trabajo y Asuntos Sociales.

[6] El Servicio Andaluz de Empleo tiene como funciones las de seleccionar los proyectos en los que quiere participar, presentando su candidatura ante la UAFSE, realizar el control financiero y su justificación y asegurar las tareas de seguimiento y evaluación.

en materia de empleo en el territorio de la Comunidad Autónoma andaluza por lo que puede utilizar la cofinanciación comunitaria para complementar los esfuerzos económicos en esta materia. Así, participa en 21 de los 29 proyectos Equal andaluces aprobados como organismo representante y cofinanciador.

Simultáneamente a la participación de la autoridad regional a través del Servicio Andaluz de Empleo y como respuesta a la multiplicidad de políticas que se ven inmersas en Equal, participan otros centros directivos (Consejerías) u organismos de la Junta de Andalucía como son las Consejerías de Medio Ambiente, Igualdad y Bienestar Social (Instituto Andaluz de Servicios Sociales, Instituto Andaluz de la Mujer, Dirección General de Bienestar Social), Gobernación (Dirección General de Coordinación de Políticas Migratorias), Innovación, Ciencia y Empresas (Dirección General de Economía Social).

— Los partenariados locales, instituciones públicas y privadas con intereses sectoriales o territoriales relacionados con los distintos colectivos beneficiaros, participan aportando su experiencia y/o cofinanciación en las Agrupaciones de Desarrollo que hay que constituir para acceder a la financiación comunitaria de esta Iniciativa.

Esta descentralización, que *a priori* parece beneficiosa puesto que se busca la adaptación de la política a cada territorio, requiere coordinación y coherencia en las actuaciones públicas para evitar duplicidad y efectos indeseados.

El estudio de los mecanismos de coordinación y comunicación y la corresponsabilidad de las distintas administraciones competentes es clave a la hora de determinar el funcionamiento de las actuaciones acometidas por cada nivel de gobierno, de acuerdo con sus competencias. Por una parte, la ausencia de canales fluidos de coordinación dificulta la correcta gestión de las medidas. Por otra parte, la evaluación resulta incompleta y poco útil a la toma de decisiones si todos los implicados no encuentran un reflejo en el proceso y sus resultados.

La ejecución y evaluación de estas políticas complejas requiere que los procesos tradicionales de coordinación se sustituyan por mecanismos y acuerdos que faciliten una gestión más activa y flexible que considere las necesidades de las organizaciones gubernamentales y otras no gubernamentales.

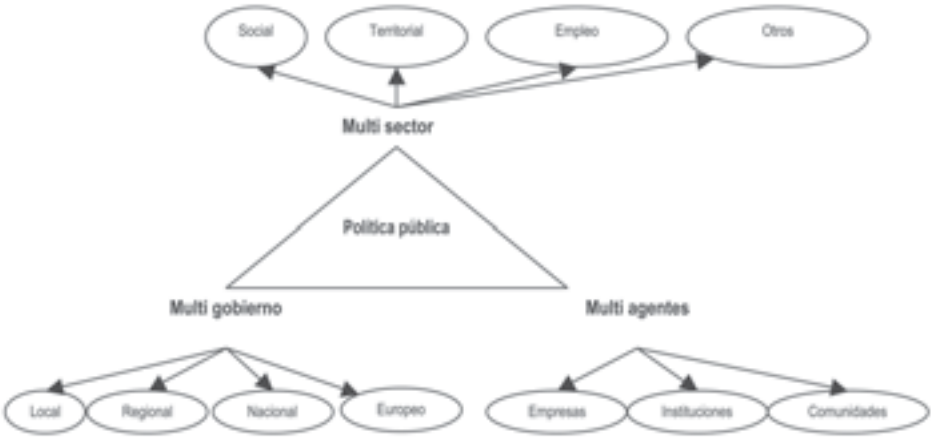
La eficacia y la legitimidad política se basa, de esta manera, y como afirma Prats (Prats, J. 2004) en «la calidad de la interacción entre los distintos niveles de gobierno y entre éstos y las organizaciones empresariales y de la sociedad civil». Por ello el diseño de la evaluación debe incorporar esta multiplicidad y heterogeneidad de actores, en un intento de facilitar el establecimiento de criterios y objetivos comunes haciendo partícipes a los implicados en el gobierno de su propio desarrollo.

6. LOS SISTEMAS DE EVALUACIÓN: UNA REALIDAD A MEJORAR

Tal y como aparece en The Guide (Comisión Europea, 2004), nos enfrentamos ahora a un nuevo reto en la determinación del enfoque metodológico que trata de reforzar las capacidades en evaluación, mediante la incorporación de un di-

Grafico n.º 2

Relación triangular compleja de la política pública



Fuente: Elaboración propia

verso grupo de instituciones y actores, tanto a escala vertical como horizontal, la consolidación de un sistema de gobernanza multi-nivel, la aplicación del principio de subsidiaridad, el establecimiento de prioridades políticas trasversales tales como el desarrollo sostenible o la igualdad de oportunidades y el reconocimiento del rol jugado por el capital social en el desarrollo económico.

A pesar de que se están realizando esfuerzos tendentes a la construcción de sistemas de evaluación que permitan recoger la complejidad de factores que intervienen en la implementación de una política, las herramientas que se están probando manifiestan algunas deficiencias y lagunas.

Por otro lado, nos encontramos con programas para los que no se prevé ningún sistema de evaluación, por incompleto que sea, bien porque aun no se crea en la bondad del medio o porque aun existe un desconocimiento de las diferencias entre evaluación y control.

En el caso del Plan de Vivienda, las autoridades responsables no habían previsto inicialmente su evaluación, ni es una exigencia reglamentaria. Sin embargo, han apostado por poner en marcha un proceso integral de evaluación como experiencia piloto en este campo que, además, es único en España para este tipo de políticas. Hasta el momento lo que se ha venido realizando, y sólo por uno de los gestores, es el seguimiento de las promociones de vi-

vienda. Ello requiere el diseño de un sistema de seguimiento y evaluación, tarea en la que nos encontramos inmersos.

Este sistema se va a estructurar en una batería de indicadores para medir y evaluar los diferentes ámbitos del Plan Municipal de Vivienda. Estos son: la ejecución y resultados de las acciones de suelo y vivienda, cobertura de la población objetivo, gestión del plan, cumplimiento de la normativa legal y evolución del contexto socioeconómico.

Los indicadores se cuantifican y valoran a la escala mínima de planificación (promociones) y posteriormente se agregan por niveles superiores (medidas, programas y plan). Así, la información se recoge siguiendo el esquema de «abajo-arriba» y los resultados se presentan de «arriba-abajo», consiguiéndose un enfoque integrador y exhaustivo del avance del plan.

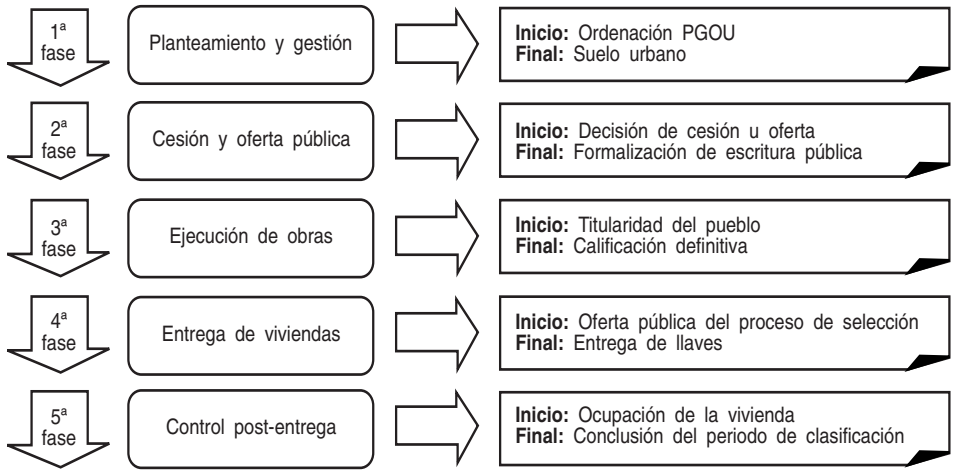
Para poder analizar la marcha del plan, se han identificado las principales fases por las que pasa una promoción de vivienda y se ha diseñado un sistema de semáforos, que en función del cumplimiento temporal de los objetivos marcados en la programación y una proyección sobre la probabilidad de realización, informará sobre la ejecución y sus posibles retardos.

Según las fases, se analiza la ejecución del plan atendiendo a las dos cuestiones siguientes:

— Datos referentes a las variables clave para evaluar el grado de avance

Gráfico n.º 3

Fases de producción de la vivienda protegida



Fuente: Elaboración propia

de las promociones: estado de los suelos y programación temporal para las acciones de suelo, y análisis por promotores, localización y características de los inmuebles, para las de vivienda.

— Estudio de los parámetros de calidad urbanística y perfil de la población solicitante y beneficiaria de las actuaciones (ámbitos de ejecución y de población objetivo). Esto es, análisis de los resultados de la ejecución del plan a partir de los niveles de medida y programa que muestran los efectos más directos.

Por su parte, la Iniciativa Comunitaria Equal, en respuesta a las exigencias de los Fondos Estructurales, debe incorporar reglamentariamente[7] un proceso de evaluación ex–ante, intermedia y ex-post (proceso desde arriba hacia abajo). El sistema de evaluación establecido traslada en cascada la responsabilidad del ámbito europeo al nacional (autoridades de gestión) y local (Agrupaciones de Desarrollo), con la intención de poder construir la evaluación en los diferentes momentos, siendo sobre el ámbito local sobre el que finalmente recae gran parte de las tareas a realizar (desde abajo hacia arriba).

De este modo, exigiendo una evaluación continua a las Agrupaciones de Desarrollo[8], que se puede gestionar de forma interna y/o externa, se persigue un proceso de mejora así como asegurar la evaluación global de la Iniciativa. Las autoridades nacionales de gestión, agregando y complementando estas evaluaciones generan los *inputs* necesarios para la evaluación a escala europea, así como las posibles reorientaciones del Programa de Iniciativa Comunitaria (PIC) en cuanto a su gestión, áreas temáticas priorizadas, colectivos beneficiarios, etc.

En Andalucía, a diferencia de otras regiones europeas, hay que considerar otro nivel, el regional, a través del Servicio Andaluz de Empleo, que ha asumido entre otras funciones, las de asegurar el seguimiento y evaluación de sus proyectos. Para el desarrollo de estas tareas ha facilitado a las Agrupaciones de Desarrollo un sistema homogéneo e informatizado de seguimiento y evaluación continuos (Sistema Equal-GSE) y ha contado con una asistencia técnica externa para apoyar estos procesos a escala regional (el Instituto de Desarrollo Regional).

Este sistema de evaluación (a escala europea, nacional, regional y local; del programa de trabajo nacional y del Acuerdo de Cooperación Transnacional-ACT) es, en última instancia, realizado por las Agrupaciones de Desarrollo que las sitúa ante la obligación de abordar procesos de evaluación para los que no

[7] El Reglamento (CE) Nº 1260/1999 del Consejo de 21 de Junio de 1999 por el que se establecen las disposiciones generales sobre los Fondos Estructurales (FFEE) establece en su título IV «Eficacia de las intervenciones de los fondos» dos capítulos dedicados al seguimiento (cap. 1, artículos 34 a 37) y evaluación (cap. 3, artículos 40 a 43), así como continuas referencias a lo largo del mismo. En el título 2 del citado reglamento se establecen las disposiciones relativas a las Iniciativas Comunitarias y a su evaluación, precisando el contenido y momento en que deben hacerse.

[8] En la Guía de Gestión y Control de Proyectos Equal (UAFSE) se establece que «(…) todos los proyectos contarán con un sistema de evaluación interna que, en su caso, podrá ser completado con la realización de una evaluación externa para garantizar su objetividad.»

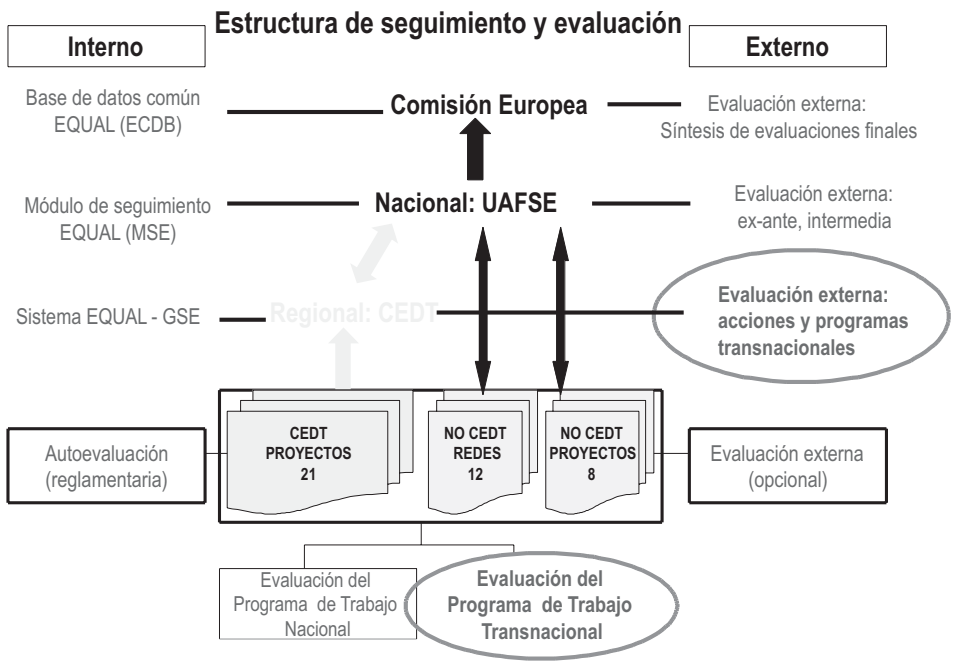
poseen la experiencia y capacidad requeridas, ya sea para realizarlos internamente o para elaborar los términos de referencia para su externalización. La incipiente comunidad evaluadora tampoco ofrece una respuesta única.

Por tanto, de la capacidad que estos partenariados desarrollen para la gestión compartida y trabajo en red, dependerá el éxito de estos programas y consecuentemente de sus evaluaciones. En Andalucía ha tenido una aplicación práctica peculiar y en el siguiente esquema se representa la estructura de seguimiento y evaluación que se ha diseñado para la Iniciativa Comunitaria Equal.

La heterogeneidad de entidades que forman parte de cada red puede ser un elemento positivo para el diseño y formulación de sistemas de seguimiento y evaluación, porque permite la especialización en diferentes etapas y en diferentes actuaciones y ámbitos. Además, pueden generarse sinergias a través de la difusión del conocimiento del funcionamiento de planes anteriores y de otras experiencias.

Gráfico n.º 4

Estructura de seguimiento y evaluación de la Iniciativa Comunitaria Equal en Andalucía



Fuente: Elaboración propia

Sumario

La existencia de aplicaciones informáticas de seguimiento y evaluación continua (como la utilizada en Equal en Andalucía) facilita y asegura la recogida de información sobre los proyectos ejecutados. No obstante, en los casos analizados, estas herramientas no han tenido la flexibilidad suficiente para recoger toda la información necesaria lo que ha requerido el diseño de otras complementarias adaptadas a las peculiaridades de cada proyecto y entidad.

En el caso de Equal, las principales debilidades del sistema de seguimiento en general han sido:

— La ausencia de indicadores especialmente significativos y explicativos sobre los aspectos horizontales de los proyectos, fundamentalmente los relacionados con el cumplimiento de los principios rectores de la Iniciativa.

— La inclusión de «campos abiertos» para la recogida de información que ha impedido la recopilación sistematizada y relevante de datos.

— La escasa capacitación para el uso de las herramientas y para el análisis e interpretación de la información a recopilar.

Cuadro n.º 2

Catalizadores de la evaluación de la Iniciativa Comunitaria Equal en Andalucía

Catalizadores	Facilitan	Frenan
Alto número de socios que forman parte de la red	Especialización Multidisciplinariedad Transferencia de conocimiento	Coordinación Dilución de responsabilidades
Experiencia previa en seguimiento y evaluación	Retroalimentación	Resistencia al cambio
Diseño de herramientas de apoyo al proceso de evaluación	Facilita el seguimiento	Deficiencias de adaptación (excesiva estandarización) Oportunidad (tiempo, coste, calidad)
Utilización de las NTIC	Facilita la recopilación de información	Excesivo encorsetamiento Diseño no orientado a la evaluación

Fuente: Elaboración propia.

Sumario

La falta de orientación de los sistemas de seguimiento hacia la evaluación, al menos tal y como se concebía inicialmente en el esquema que se ha mostrado, es uno de los motivos que ha restado éxito al proceso evaluador de Equal.

Esto es, la necesidad de dar cumplimiento a diferentes exigencias de evaluación (para la Comisión, para la autoridad nacional, para el gobierno regional y para las Agrupaciones de Desarrollo) ha causado desconcierto en los centros directivos encargados de recoger la información, que han debido aportar los mismos *inputs* en diferentes ocasiones y a diferentes equipos de evaluadores. Esto ha llevado a un agotamiento de las fuentes de información y a que las Agrupaciones de Desarrollo centraran el proceso en el seguimiento y no en la evaluación, que además de continua, debería haberse realizado con participación interna de los técnicos de los proyectos.

Y como muestra, sólo un 29% de las evaluaciones realizadas en Equal Andalucía han sido internas, y en cada una sólo ha participado directamente uno de los socios de las agrupaciones manteniendo los demás un papel consultivo. Ello ha limitado la capacidad de integración con el proyecto y la institucionalización de la evaluación continua que se perseguía.

Teniendo en cuenta la escasa cultura evaluativa de las instituciones y la débil capacidad para realizarlas, la propuesta de basar el proceso en una autoevaluación continua a nivel de Agrupaciones de Desarrollo no parece ayudar a su desarrollo e institucionalización.

Si a esto se le une el retraso en el inicio de las evaluaciones, se puede concluir que el sistema de evaluación que se ha llevado a la práctica no ha sido de gran utilidad para la toma de decisiones, puesto que las recomendaciones, cuando las ha habido, no han podido incorporarse a tiempo.

7. DE LA DESTILACIÓN EN CASCADA A LA GENERACIÓN DE UN SISTEMA DE EVALUACIÓN

En definitiva, tal y como hemos ido analizando, hay una serie de aspectos que condicionan los procesos evaluativos que se pueden poner en marcha y el propio diseño del sistema de evaluación para que sea útil y eficaz. Entre ellos destacan los siguientes:

- El amplio partenariado de socios (incluso transnacionales) que participan en la gestión de los proyectos entre los que se identifican administraciones públicas multinivel con agentes sociales y entidades del tercer sector muy heterogéneas con intereses diferentes a considerar.
- La concurrencia de múltiples políticas horizontales y sectoriales: empleo, social, igualdad de oportunidades, medio ambiente, sociedad de la información, etc.
- La preocupación por la utilidad y utilización de la evaluación para incorporar a las políticas generales las lecciones innovadoras.
- El carácter experimental de programas que persiguen desarrollar la capacidad de innovación en la elaboración, gestión y evaluación de las políticas.

Esta estructura requiere una nueva aproximación metodológica que contribuya a crear un sistema de evaluación arti-

culado y coherente que permita valorar las siguientes cuestiones:

— El alineamiento y coherencia de los diferentes objetivos perseguidos por las distintas políticas sectoriales y horizontales y el impacto de estas actuaciones en la población beneficiaria.

  Porque los programas no se diseñan ni ejecutan aislados. Forman parte de un mosaico donde cada pieza juega un papel en los resultados finales de un proceso socio político más amplio y complejo.

— El diseño y desarrollo de un sistema de seguimiento capaz de satisfacer las necesidades de información de todas las instituciones y agentes implicados, y que considere la dimensión dinámica de las políticas y de las redes. Se concibe así a los programas como procesos, más que productos, es decir, como un conjunto abierto y dinámico, donde debe contemplarse la evolución de objetivos e intereses de los actores que intervienen en su gestión y ejecución, produciendo información útil para la negociación de nuevas actuaciones.

  Además, la evaluación debe proporcionar información para generar conocimiento. Ello fomenta la utilización de la evaluación por las instituciones, que no ven la evaluación como un simple control final de resultados sino como un análisis de una realidad más compleja que pretende ayudarles a ser mejores.

— El establecimiento de criterios de evaluación que considere las expectativas y necesidades de los implicados en los programas.

— La adaptación del enfoque del análisis de redes sociales para determinar la relación entre los actores sociales, la estructura de la red que configuran y su funcionamiento. Para su aplicación se debe estudiar la centralidad de las posiciones de los actores, los círculos sociales, las subredes, en definitiva, las relaciones que mantienen los actores más que a cada uno de ellos considerados individualmente.

  La pertinencia de este enfoque reticular para conocer el proceso de participación y capacitación, se fundamenta en el presupuesto de que las estructuras creadas son un condicionante para la formación de interés colectivo, la definición de objetivos comunes, la constitución de organizaciones y alianzas y su sostenibilidad.

  Este enfoque, al centrarse en las relaciones establecidas entre los distintos nodos de la red, y no en los sujetos que las mantienen, nos aproxima a la medición o delimitación del capital social, como recurso empleado/generado en el proceso de compartir información y conocimiento al proceso de elaboración de políticas.

Ello puede implicar la puesta en marcha de un modelo descentralizado donde la responsabilidad de la evaluación recae en los gestores de los programas.

El objetivo es fortalecer las capacidades de las organizaciones, aplicando técnicas de evaluación en el proceso de elaboración de políticas públicas, contribuyendo así a la potenciación de las instituciones y al proceso de aprendizaje institucional.

Sumario

REFERENCIAS BIBLIOGRÁFICAS

AGUILAR, M.J. Y ANDER-EGG, E. (1992): *Evaluación de Servicios y Programas Sociales.* Siglo XXI Madrid.

AWASTHI, D.N. Y SEBASTIAN, J. (1996): *Evaluation of entrepreneurship development programmes.* Sage Publications. Delhi, India.

BAÑÓN I MARTÍNEZ, R. (Compilador) (2003): *La evaluación de la acción y las políticas públicas.* Díaz de Santos. Madrid.

BERK, R.A. Y ROSSI, P.H. (1999): *Thinking about Program Evaluation 2.* Sage Publications. EEUU.

BLACK, T.R. (1998): *Evaluating Social Sicence Research. An introduction.* Sage Publications. Trowbrigde, Wiltshire. RU.

BOYLE, R. Y LEMAIRE, D. (1999): *Building Effective Evaluation Capacity: Lessons from practice.* Transaction Publishers. New Brunswick, New Jersey. EEUU.

BUSTELO RUESTA, M. (2003): *Evaluation of gender mainstreaming: ideas from a meta-evaluation. Evaluation 4.* Sage. Londres. UK.

CLARKE, A. (1999): *Evaluation Research. An introduction to principles, methods and practice.* Sage Publications. Trowbridge, Wiltshire.RU.

COHEN, E. Y FRANCO, R. (1993): *Evaluación de proyectos sociales.* Siglo Veintiuno de España. Madrid.

COMISIÓN EUROPEA (1999): *MEANS (6 volúmenes).* Oficina para las publicaciones oficiales para la UE. Luxemburgo.

COMISIÓN EUROPEA (2004): *The Guide*. Publication webside: http://www.evalsed.info/

COOK, T.D. Y REICHARDT, CH.S. (1986): *Métodos cualitativos y cuantitativos en investigación evaluativa*. Ediciones Morata. Madrid.

CRONBACH, L.J. (1982): *Designing evaluation of educational and social programs*. Jossey-Bass. San Francisco. EEUU.

DALE, R. (1998): *Evaluation Frameworks for Develpment and Projects.* Sage Publications. Nueva Delhi. India.

DROR, Y. (1974): *Public policy making reexamine.* Leonard Hill Books. Nueva York. EEUU.

EPSTEIN, I., TRIPODI, T. (1977): *Reaserch techniques for programme planning, monitoring and evaluation.* Columbia University Press. Nueva York. EEUU.

EVERITT, A. Y HARDIKER, P. (1996): *Evaluating for good practice.* Macmillan Press. Malasia.

FERNÁNDEZ BALLESTEROS, R. (Ed.) (1995): *Evaluación de programas. Una guía práctica en ámbitos sociales, edicativos y de salud.* Síntesis Psicología. Madrid.

GALTUNG, J. (1966): *Teoría y métodos de la investigación social.* Endeba. Buenos Aires. Argentina.

GUBA, E.G. Y LINCOLN, Y. S. (1989): *Fourth Generation Evaluation.* Sage Publications. EEUU.

HAKIM, C. (1988): *Research design.* Allen and Unwin. Londres. RU.

HEMMATI, M. (2002): *Multi-stakeholder Process for Governance and Sustainability.* Erathscan Publications Ltd. London.

HOUSE, E.R. (1993): *Professional Evaluation. Social impact and political consequences.* Sage Publications. EEUU.

HOUSE, H. (1973): *The politics and process.* McCutchan. Nueva York. EEUU.

INSCAL (2001): *Guía de autoevaluación para la Administración Pública: modelo EFQM de excelencia*. Ministerio de Administraciones Públicas. Ministerio de la Presidencia.

INSTITUTO ANDALUZ DE LA MUJER (2001): *Mainstreaming o Enfoque Integrado de Género: Manual de aplicación en proyectos de empleo*. Likadi. Madrid.

INSTITUTO DE DESARROLLO REGIONAL Y SERVICIO ANDALUZ DE EMPLEO (2005): *La Iniciativa Comunitaria Equal. Actuaciones en Andalucía: Informe final 2001-2004*. Dirección General de Fomento del Empleo. Servicio Andaluz de Empleo. Consejería de Empleo. Junta de Andalucía. Sevilla.

INSTITUTO DE DESARROLLO REGIONAL Y SERVICIO ANDALUZ DE EMPLEO (2005): *La Iniciativa Comunitaria Equal. Actuaciones transnacionales: Informe final 2001-2004*. Dirección General de Fomento del Empleo. Servicio Andaluz de Empleo. Consejería de Empleo. Junta de Andalucía. Sevilla.

Sumario

INSTITUTO NACIONAL DE ADMINISTRACIONES PÚBLICAS (2000): *Evaluación y Calidad en las Organizaciones Públicas.* Ministerio de Administraciones Públicas. Madrid.

JONES, C. (1977): *An introduction to the study of public policy.* Duxbury. Boston. EEUU.

KEVIN WILLIAMS, K.; DE LAAT, B. Y STERN E. (2002): *The Use of Evaluation in the Commission Services*.

LINDBLOM, C Y WOOD-HOUSE, E.J. (1991): *The policy making process*. Prentice Hall. Englewood Cliffs, NJ. EEUU.

LÓPEZ CASASNOVAS, G. (2003): *Los nuevos instrumentos de la gestión pública.* La Caixa. Barcelona.

MENY, I. Y THOENING, J.C. (1992): *Las políticas públicas.* Ariel Ciencia Política. Barcelona.

MOHR, L.B. (1995): *Impact analysis for program evaluation.* Sage Publications. EEUU.

MONNIER, E. (1995): *Evaluación de la acción de los poderes públicos.* Instituto de Estudios Fiscales. Ministerio de Economía y Hacienda. Madrid.

OCDE (1992): *La evaluación de programas de empleo y de medidas sociales. Estado actual de una cuestión compleja.* Ministerio de Trabajo y Seguridad Social. Madrid.

OCDE PROCEEDINGS (1997): *Policy Evaluation in Innovation and Technology: towards best practices*.

OGANDO CANABAL Y MIRANDA ESCOLAR (COORD.) (2003): *Evaluación de programas e iniciativas comunitarias: experiencias, nuevas orientaciones y buenas prácticas*. Instituto de Estudios Europeos. Valladolid.

OSUNA LLANEZA, J. L. Y MÁRQUEZ, C. (COORD.) (2000): *Guía para la evaluación de las políticas públicas.* IDR. Sevilla.

OSUNA LLANEZA, J. L. Y VÉLEZ MÉNDEZ, C. (2004): *Guía de Planificación y Evaluación para Agentes de Desarrollo Local.* Sevilla Siglo XXI, S.A. Sevilla.

OWEN, J.M. (1999): *Program Evaluation. Forms and approaches*. Sage Publications. Singapur.

PATTON, M.Q. (1997): *Utilization Focused Evaluation. The New Century Text*. Sage Publications. EEUU.

PAWSON, R. Y TILLEY N. (1997): *Realistic Evaluation.* Wiltshire. RU.

PEARCE, S.W. Y NASH C.A. (1981): *The social appraisal of projects.* MacMillan. Londres. RU.

PLANTE, J. (1991): *En plan-ENA: outlis, practiques et institutions pour èvaluer les politiques publiques.* La Documentation Française. Paris. Francia.

POSAVAC, E.J. (1980): *Program evaluation.* Prentice-Hall. Englewood Cliffs, NJ. EEUU.

PRATS, J. (2004): «Las transformaciones de las administraciones publicas de nuestro tiempo» en *Instituciones y Desarrollo* [http://www.iigov.org; 16/07/2004].

PRIOR, D.; VERGES, J. Y VIUDALRDELL, I. (1993): *La evaluación de la eficiencia de los sectores públicos y privados*. Instituto de Estudios Fiscales. Madrid.

REBOLLO CATALÁN, M.A., COLÁS BRAVO, M.P. (2002): *Evaluación de programas: una guía práctica*. Kronos. Sevilla.

ROSSI, P.H., FREEMAN, H.E. Y LIPSEY, M. (2003): *Evaluation: a systematic approach.* 7ª edición. Sage Publications. EEUU.

SABATIER, P.A. Y JENKINS-SMITH, H.C. (EDIT.) (1993): *Policy Change and Learning. An advocacy colation approach*. Westview Press. Colorado. EEUU.

SANTOYO, H., RAMÍREZ, P. Y SUVEDI, M. (2002): *Manual para la evaluación de programas de desarrollo rural.* Mundi-Prensa. México.

SHAW, I.F. (1999): *Qualitative Evaluation.* Sage Publications. Trowbridge, Wiltshire. RU.

SHAW, I.F. Y LISHMAN J. (Ed.). (1999): *Evaluation and Social Work Practice*. Sage Publications. Trowbridge, Wiltshire. RU.

STUFFLEBEAM, D.L.; SHIRKFIELD, A.J. (1987): *Evaluación sistemática: guía teórica y práctica.* Paidós. Barcelona.

SUBIRATS, J. (1994): *Análisis de políticas públicas y eficacia de la Administración.* Ministerio de Administraciones Públicas. Madrid.

SUCHMAN, E. (1967): *Evaluation reaserch*. Russel Sage. Nueva York. EEUU.

TAYLOR, S.J.; BOGDAN, R. (1986): *Introducción a los métodos cualitativos de investigación.* Paidós. Buenos Aires. Argentina.

VÁZQUEZ BARQUERO, A. (1999): *Desarrollo, redes e innovación.* Pirámide. Madrid.

Sumario

VEDUNG, E. (2000): *Public policy and program evaluation.* Transaction Publishers. New Brunswick. NJ. EEUU.

VÉLEZ MÉNDEZ, C. (2000): «La evaluación de políticas públicas en Andalucía: una realidad en proceso de gestación», en *Cuadernos Económicos de Granada nº 11: Estrategia y tácticas de desarrollo regional y convergencia económica: urbanismo, coyuntura inmobiliaria y sistema financiero.*

WEISS, H. C. (1990): *Investigación evaluativa: métodos para determinar la eficiencia de los programas de acción.* Trillas. México.

WEISS, H. C. (1998): *Evaluation.* Prentice-Hall. Upper Saddle River, NJ. EEUU.

WHOLEY, J.S. (1979): *Evaluation: promise and performance.* Urban Institute. Washington DC. EEUU.

WHOLEY, J.S. (1983): *Evaluation and effective public management.* Little Brown. Boston. EEUU.

Sumario

Principales mecanismos de evaluación económica de políticas públicas

El diseño e implantación de políticas públicas requiere un proceso de evaluación que debe incluir la consideración del coste y de los beneficios de las mismas. Este artículo describe ordenadamente las distintas metodologías y técnicas utilizadas en estos análisis, centrándose especialmente en los métodos cuantitativos más apropiados para el propósito en cuestión. La revisión de la literatura que hemos realizado sirve también para poder apreciar el distinto desarrollo de la aplicación de cada una de estas técnicas en nuestro país. La evaluación económica del Análisis Coste-beneficio y similares, las Fronteras Estocásticas y el Análisis Envolvente de Datos aparecen como las técnicas microeconómicas más aplicadas, mientras que entre las técnicas macroeconómicas destacan las Tablas input-output.

Politika publikoak diseinatu eta ezartzeak ebaluazio-prozesua eskatu egiten du eta prozesu horretan gainera politika publikoen kostuaren eta mozkinen azterketa sartu behar da. Artikuluan analisi hauetan erabiltzen diren metodologia eta teknikak deskribatu dira ordena handiz, eta delako xedea lortzeko aproposenak diren metodo kuantitatiboetan batik bat arreta jarri da. Literaturaz egin dugun berrikuspenak teknika hauen aplikazioak gure herrialdean izan duen garapenaz konturatzeko balio du. Kostuaren eta mozkinen (eta antzerakoen) Analisiaren ekonomia-ebaluazioa, Muga Estokastikoak, eta Datuen Analisi Inguratzailea dira, hain zuzen ere, mikroekonomiaren teknikarik erabilienak; makroekonomiaren tekniketan, berriz, input-output taulak dira behinenak.

The design and implementation of public policies is needed of an evaluation process, which should take into account their cost and benefits. This article summarises in an ordered way the different methodologies one could use for these analysis, mainly focusing in quantitative techniques. The literature review shows the different development that these techniques in our country. Cost-benefit Analysis, Stochastic Frontiers and the Data Envelopment Analysis seem to be the most applied microeconomic techniques, while among the macroeconomic ones input-output tables points up.

Ivan Planas Miret
Universitat Pompeu Fabra

INDICE

1. Introducción. Mecanismos de evaluación de políticas y reformas sociales
2. Mecanismos y técnicas de evaluación económica de políticas y reformas sociales
3. A modo de conclusión. Los problemas más usuales en la evaluación económica y algunas recomendaciones
Referencias bibliográficas
Anexo

Palabras clave: gestión pública, evaluación económica, eficiencia

N. de clasificación JEL: H21, H22, H43

1. INTRODUCCIÓN. MECANISMOS DE EVALUACIÓN DE POLÍTICAS Y REFORMAS SOCIALES

Tal y como señala la exposición de motivos de la Ley 6/1997, de 14 de abril de 1997, de Organización y Funcionamiento de la Administración General del Estado, «La Constitución de 1978 (RCL 1978, 2836) ha supuesto una profunda alteración del sistema de fuentes del Derecho público de manera que la actividad y la estructura de la Administración General del Estado se encuentran vinculadas por el marco constitucional. El precepto fundamental que dedica nuestra vigente Constitución a la Administración General del Estado es el artículo 103 en el que se recogen los principios básicos que deben presidir la actividad de la Administración estatal, a saber: servicio, objetividad, generalidad, eficacia, jerarquía, descentralización, desconcentración y coordinación.»

Recuerda, además, la ley mencionada en su Artículo 3, «La Administración General del Estado se organiza y actúa, con pleno respeto al principio de legalidad, y de acuerdo con los otros principios que a continuación se mencionan:[...] 2. De funcionamiento. a) Eficacia en el cumplimiento de los objetivos fijados. b) Eficiencia en la asignación y utilización de los recursos públicos.»

Es decir, el cambio de mentalidad de la administración hacia la gestión pública conlleva, también, la exigencia ética y constitucional de la eficiencia en el sector público. Dicha exigencia requiere, por

tanto, un ejercicio de validación de los resultados obtenidos de la acción pública y, por consiguiente, el desarrollo de técnicas que permitan su evaluación.

Cuando se habla de métodos de evaluación de políticas públicas, se engloba un amplio conjunto de técnicas que se aplican dependiendo del objetivo del análisis o de la política que se desee seguir, del momento en que se ha llevado a cabo dicho análisis (previo a la toma de decisiones o posterior a ésta), de la singularidad de la medida (susceptible, por ejemplo, de prueba piloto...), e incluso del sector económico sobre el que se aplique (sanidad, obras públicas, educación..). Determinadas técnicas requieren una amplia disponibilidad de datos y sus consiguientes valoraciones económicas (como, por ejemplo, el Análisis coste-beneficio), otras requieren el computo de variables externas al proyecto (como la medición de los efectos distributivos de una determinada política), y, finalmente, otras pueden aplicarse midiendo el rendimiento sobre la unidad de análisis (tales como la efectividad o la eficacia en la consecución de resultados).

Como en cualquier ejercicio de análisis, la evaluación de proyectos y políticas públicas se divide en dos grandes áreas: la del análisis cualitativo y la del análisis cuantitativo. La Economía, y así la evaluación económica, se ha especializado principalmente en el análisis cuantitativo.

Hay que indicar que los distintos métodos de evaluación económica de proyectos públicos no han sido desarrollados, sin embargo, por una sola disciplina, sino por distintas especialidades que han trabajado de forma independiente en el desarrollo de estos instrumentos (ciencias del medio ambiente, de la salud, economía del bienestar, e incluso fuera de la propia Economía en áreas interdisciplinares como la gestión pública o las propias ciencias políticas).

Los objetivos por los que se realizan las evaluaciones son diversos y, sin ánimo de ser exhaustivos, algunos se exponen a continuación: a) demostrar que el programa elegido y los mecanismos a través de los cuales se ha implantado son acertados; b) determinar si un programa funciona o no adecuadamente; c) determinar su eficacia en la consecución de los objetivos inicialmente pretendidos; d) justificación de gastos presupuestados o ya ejecutados; e) apoyar la extensión de un programa o su generalización a distintas zonas del país; f) justificación ante los directores del programa de la bondad de la iniciativa; g) retrasar la introducción o puesta en marcha de un nuevo proyecto, en razón precisamente de la petición de pruebas de efectividad o eficiencia que lo justifiquen.

1.1. **Los diferentes mecanismos, un poco de historia**

Los mecanismos de control (y evaluación) de políticas han evolucionado históricamente siendo los principales hitos:

a) la auditoría: constata la veracidad en el cumplimiento de los objetivos de la gestión de gobierno, sin cuestionar su validez, dimensión, o adecuación. Esencialmente, se basa en la verificación documental, e incluye una dimensión de la eficacia / efectividad de la política pública;

Sumario

b) el control de gestión: desarrollado en los años sesenta del siglo XX como resultado de la planificación de la acción de gobierno. Consiste en analizar la correspondencia entre los planes de gobierno y las acciones realizadas;

c) el presupuesto por programas: técnica de presupuestación (originada en EEUU y Francia, en la década de 1970) que relaciona el uso de recursos con los objetivos presupuestarios. Es la base del actual análisis de costes;

d) el presupuesto por resultados, ligado a la tecnología informática, que permite desarrollar grandes series de datos, construir indicadores para el seguimiento de las actividades del Gobierno, y verificar así el cumplimiento de las políticas;

e) el gerencialismo público, iniciado en la década de 1960, promovida por las Naciones Unidas, que hace hincapié en la utilización generalizada de indicadores de gestión;

f) la evaluación de programas, ampliamente vinculada al desarrollo del Estado de bienestar (programas con objetivos concretos y poblaciones objetivo específicas).

Aunque para proyectos concretos y de forma aislada sí se utilizan estos diversos métodos, todavía no han sido incorporados de forma sistemática a la gestión de la administración pública.

Como señala Emilio Albi (2000), «la evaluación de la eficiencia pública y la mejora de la gestión no se improvisan. Se pueden realizar evaluaciones ad hoc, aisladas, de determinados programas públicos, pero en este caso es previsible encontrar la resistencia lógica de los responsables de estos programas que no quieren ser los únicos fiscalizados. Obtener información estadística y contable necesaria requiere planificación, tiempo y tenacidad. La formación del personal en las técnicas de evaluación es, igualmente, un asunto difícil y costoso. Si se quiere pues controles de eficiencia, hay que establecer la obligación legal, aplicando la voluntad política y decidiendo quién va a llevarla a cabo».

Los diferentes mecanismos desarrollados pueden ser utilizados para la evaluación prospectiva de las políticas, la supervisión de su evolución y el análisis final de resultados. Mientras la definición del análisis de políticas se ha basado más en la idea prospectiva, la evaluación de programas o de políticas públicas se sustenta más en la idea de revisión retrospectiva con el objetivo de estimar los efectos de actuaciones futuras (Poland, 1974). Según Chelimsky (1985), y tal y como recoge Subirats (1994), el gran desarrollo de la evaluación de programas se produce en la década de 1970, debido principalmente a los siguientes factores: las demandas del Congreso de los EEUU de información procedente del proceso de aplicación de las leyes por parte del ejecutivo, la demanda de conocimiento general sobre unas políticas que precisaban de un gran esfuerzo inversor, y la necesidad de los gestores públicos y altos cargos de la Administración de conocer el estado de aplicación de los programas en curso y sus efectos.

Sumario

1.2. El desarrollo de la evaluación de políticas públicas en nuestro país

La reciente creación, por parte del Ministerio de Administraciones Públicas,[1] de la Agencia Estatal de Evaluación de Políticas Públicas y Calidad de los Servicios abre una puerta a la normalización de la aplicación de los mecanismos de evaluación en las diversas administraciones públicas de nuestro país. La implantación generalizada de la evaluación conlleva muchas ventajas, tales como, la de forzar la recogida sistemática de datos que permita mejorar el ajuste de los resultados a las distintas circunstancias, la posibilidad de generar estudios de metanálisis o estudios comparativos de políticas aplicadas en distintos puntos del país y/en ámbitos similares, la creación de un cuerpo de técnicos y gestores con conocimientos específicos en esta área, etc. En definitiva, permite multiplicar los efectos de los actuales estudios que se realizan de forma discrecional y aleatoria.

Actualmente, las funciones de la Agencia estatal son, más que de tipo ejecutivo, de promoción de la cultura de la evaluación (impulsarla, proponer la implantación de indicadores, elaborar guías y metodologías), aunque dos objetivos parecen destacar en cuanto a su protagonismo futuro: el de desarrollar un marco normativo y metodológico adecuado para promover la calidad de los servicios públicos, y el de preparar, junto con otras instituciones y administraciones, el Plan Bienal de Evaluaciones que deberá ser aprobado por el Consejo de Ministros.

Esta Agencia estatal puede, según lo expuesto, representar un cambio en la cultura de la Evaluación en nuestro país. En este sentido, cabe esperar que el resto de administraciones públicas, especialmente las autonómicas, creen coordinadamente también sus propios órganos homólogos con similares propósitos que no puede sino redundar en la mejora de la eficiencia pública.

2. MECANISMOS Y TÉCNICAS DE EVALUACIÓN ECONÓMICA DE POLÍTICAS Y REFORMAS SOCIALES

Intentaremos en esta sección trazar un mapa de los distintos métodos de evaluación económica disponibles en la actualidad. En este artículo se abordarán principalmente los mecanismos de evaluación cuantitativa de las políticas, dejando sin explorar los mecanismos cualitativos (entrevistas con profundidad, métodos observacionales, revisión documental) más usuales en determinados ámbitos de evaluación como el educativo (13, 40), y que exigen juicios de valor más explícitos. Ello no implica que los estudios cualitativos no tengan una gran importancia; de hecho, algunos autores abogan por integrar ambos tipos de estudios, analizando primero las observaciones de forma cualitativa (individuos participantes y no participantes, cambios en el entorno, evolución de condicionantes, entre otras), para aplicar después las técnicas cuantitativas con el fin de concretar, si es posible, las observaciones cualitativas. La economía de la salud tiene, a este respecto, una extensa literatura sobre diferentes aproximaciones a la cuantificación de aspectos cualitativos de los resultados (calidad de

[1] Orden Ministerial 3017/2004, de 16 de septiembre, a raíz de las reformas legislativas previas: Orden Ministerial 1463/2004, de 19 de mayo ; Real Decreto 1320/2004, de 28 de mayo

vida, dolor, movilidad, etc.), así como la tiene la economía del medio ambiente en lo que atañe al impacto visual, molestias, impacto sonoro, etc., de las actuaciones evaluadas.

Para poder centrarnos en las distintas clasificaciones que se desarrollarán a continuación, deben distinguirse los distintos procesos que implica la realización de la evaluación de un proyecto o de una política pública.

El proceso de evaluación requiere, inicialmente, un diseño, es decir, especificar cómo distinguir entre la nueva situación (del sector, del territorio, del grupo de población, por ejemplo), y aquella en la que se estaría de no haberse aplicado la política (escenario conocido como «contrafactual»). La existencia o no de este contrafactual, o la generación del mismo de forma artificial, nos lleva a distinguir entre diseños experimentales, cuasiexperimentales, y no experimentales. Todos los métodos que describiremos a continuación entrarían en alguna de estas categorías.

En el diseño experimental, el evaluador organiza de manera aleatoria la distribución de los individuos que deciden participar en el programa, ya sea como grupo de tratamiento, o como grupo control (este es el diseño tradicional de los ensayos clínicos experimentales o aleatorios). A pesar de ser uno de los métodos más sólidos, su utilización presenta algunos problemas: éticos, en la asignación de los grupos de afectados; dificultad para realizar pruebas piloto sin extender el programa a toda la población; coste de los experimentos y control efectivo para que la asignación sea realmente aleatoria.

Ahora bien, en muchos casos, el principal problema es que la no estandarización de los procesos de evaluación de políticas lleva a que se planteen precisamente después de que la política haya sido llevada a cabo. Las dificultades éticas en la asignación aleatoria de las pruebas piloto se resuelve, comúnmente, mediante la elección de los individuos participantes según criterios externos. La evaluación de políticas públicas mediante métodos experimentales es, en todo caso, más propia de países como EEUU que el nuestro, donde las dificultades empíricas ocasionan que la mayoría de estudios adopten enfoques de tipo cuasi-experimental o no experimental.

Los métodos de diseño cuasiexperimental consisten en construir de forma artificial el grupo de control. Las técnicas estadísticas generalmente asociadas a estos diseños (*matching*, *score matching*, comparación reflexiva) serán descritas dentro de las clasificaciones que realizaremos a continuación. Finalmente, las técnicas aplicadas a los diseños no experimentales (como la técnica de variables instrumentales o el análisis de regresión multivariante) tienen la ventaja de que resultan más sencillas y de coste reducido, aunque su solidez estadística es menor.

2.1. **Análisis microeconómicos de evaluación de políticas públicas**

En una primera clasificación, deben distinguirse las siguientes categorías:

2.1.1. Mecanismos según el objetivo del análisis

Si clasificamos los tipos de evaluación según el objetivo pretendido, hablamos entonces de: 1) Determinación de necesida-

des, 2) Evaluación formativa o correctora, y 3) Evaluación de balance o conclusiva.

La evaluación económica de proyectos públicos no se realiza siempre con los mismos objetivos. El análisis puede llevarse a cabo con la intención de determinar las necesidades de recursos (financieros, materiales...) en la fase de estudio de la viabilidad de un proyecto o, en caso de partir ya de su aprobación pública, antes de iniciar la gestión del mismo. En otros casos, el objetivo del análisis se basa en la corrección de ineficiencias surgidas en el proceso de ejecución, y por tanto el estudio se centra en aquellos recursos y costes gestionables o evitables. Y, finalmente, en algunos casos el objetivo del análisis es realizar un informe final de balance que permita conclusiones bien fundamentadas.

2.1.2. Mecanismos según el indicador de gestión

En un primer grupo podemos encontrar aquellos mecanismos más simples que consisten en la comparación de indicadores de gestión, bien sean los costes, la eficacia, la efectividad, la eficiencia, la equidad y/o los beneficios. Así, si se toma la efectividad como indicador de referencia, se entendería como preferible aquel proyecto público cuya efectividad (nivel de consecución de los objetivos pretendidos por cada proyecto alternativo dado un mismo nivel de recursos) sea superior. En otros casos el indicador de gestión puede basarse simplemente en el resultado, sin la previa elaboración de ningún índice de eficacia, efectividad..., así podría considerarse la comparación entre tasas de éxito, tasas de fallos por unidad de coste...

Todos estos indicadores tienen como objetivo, o bien el análisis de coste, o bien el análisis de resultados. En un segundo grupo que desarrollaremos más adelante en el texto se analizarán los indicadores compuestos cuyo objetivo consiste en el análisis de la comparativa entre coste y resultados.

Los índices de eficacia, efectividad, y eficiencia suelen exponerse como indicadores de calidad del proceso. Por eficacia se entiende, en el caso de un servicio, la probabilidad de que un individuo se beneficie de una tecnología o proceso bajo condiciones de actuación ideales (validez universal). Por efectividad se entiende el mismo concepto, aunque bajo condiciones reales de actuación (entorno real con los condicionantes de actuación que suceden en un determinado periodo).

Finalmente, se distinguen tres tipos de eficiencia, clasificación iniciada por Farrell MJ, (1957), quien describe también las bases del Análisis Envolvente de Datos desarrollado por Charnes, Cooper y Rodees (1978): la técnica (maximización de resultado con un determinado volumen de recursos), la productiva (maximización de resultado dados unos costes determinados, o un precio determinado de los recursos), y, finalmente, la económica (producción de los bienes y servicios que aportan más valor a la sociedad, al menor coste social posible). Así, los tres principales problemas relativos a la calidad son: la sub-utilización, la sobre-utilización y la mala utilización. Se habla también de eficiencia a escala cuando una unidad de producción es completamente eficiente pero puede aumentar su productividad incrementando la escala o nivel de producción.

El siguiente cuadro recoge los distintos tipos de eficiencia y sus métodos de estimación:

Sumario

Cuadro n.º 1

Métodos de estimación de la eficiencia en la producción de bienes y servicios.

Métodos	Tipo de eficiencia que mide	Estimación funcional	Función distribución	Ventajas
Método de comportamiento medio esperado	Global	Paramétrica	Función de costes	No es necesario determinar la frontera – coste mínimo posible Su especificación funcional es bastante simple Permite óptimos locales
Modelo de frontera determinista	Global	Paramétrica	Función de costes	No precisa datos de carácter temporal Permite óptimos locales
Modelo de frontera determinista	Técnica	Paramétrica	Función de producción	No precisa datos de carácter temporal Permite óptimos locales
Modelo de frontera estocástica	Global	Paramétrica	Función de costes	Permite captar el efecto de alteraciones externas sobre el resultado Permite óptimos locales
Modelo de frontera estocástica	Técnica	Paramétrica	Función de producción	Permite captar el efecto de alteraciones externas sobre el resultado Permite óptimos locales
Análisis Envolvente de Datos	Técnica	No paramétrica	Sin forma funcional (utilizan información de consumos y producción)	Simplicidad No es necesario hacer ninguna suposición sobre la forma funcional de la función de producción. No requiere tantas observaciones como el FDH.
FDH (Free Disposal Hull)	Técnica	No paramétrica	Sin forma funcional (utilizan información de consumos y producción	No es necesario hacer ningún supuesto sobre la forma funcional de la función de producción.
Funciones de Distancia	Asignativa	No paramétrica	Sin forma funcional (utilizan información de costes y producción	

Fuente: Elaboración propia

De todas estas técnicas desarrolladas durante los últimos 40 años, los dos métodos más utilizados son el Análisis Envolvente de Datos (DEA) y las Fronteras Estocásticas, tal como puede comprobarse en la tabla 3 al ver el número de publi-

caciones llevadas a cabo en nuestro país sobre cada una de estas técnicas.

Dadas las anteriores definiciones, la evaluación de proyectos consiste en la medición de la magnitud de estos indicadores. La medición de eficacia (al ser un indicador de resultado en condiciones ideales) quedaría relegada a la fase previa a la introducción del programa. La medición de efectividad y eficiencia servirían para determinar si la elección de recursos y de proceso productivo son los adecuados, y si existen posibilidades de mejora.

La medición de la eficiencia técnica requiere la utilización de fronteras de mejora práctica (o bien real —cuando tenemos en cuenta las unidades productivas más eficientes—, o bien estimada, cuando se determina a partir de un modelo teórico de producción). Las técnicas de medición incorporan conceptos de programación lineal y técnicas econométricas.

En un segundo grupo encontraríamos aquellos mecanismos basados en la comparación de dos tipos de indicadores: a) análisis coste-eficacia, b) análisis coste-efectividad, c) análisis coste-beneficio, y d) análisis coste-utilidad[2].

A diferencia de los anteriores, aquí lo relevante no es sólo la magnitud de cada indicador (según el cual diríamos que un proyecto A es más efectivo, más eficiente que otro), sino que se analiza cada proyecto en términos relativos.

El *Análisis Coste-eficacia* es un método de comparación de proyectos en condiciones ideales o no ambientales (condiciones de laboratorio). Este tipo de análisis es el que se suele realizar de manera previa a la introducción de un nuevo programa. A partir de alternativas más o menos parecidas y considerando la eficacia (efectividad esperada) de los recursos, se obtiene este indicador que posteriormente se podrá confrontar con el del análisis coste-efectividad, si así se desea. No es extraño que determinados proyectos puedan tener una ratio positiva coste-eficacia, y en el momento de su implantación se compruebe que la distancia entre la eficacia y la efectividad (por múltiples motivos) es tan grande que la ratio coste-efectividad es negativa (véase, a este respecto, el artículo de Ortún, en este mismo monográfico).

El *Análisis Coste-efectividad* es un método para la comparación de proyectos o políticas alternativos, así como de programas o unidades en competencia. En este análisis, los costes se miden en unidades monetarias, y los resultados (efectividad) en unidades no monetarias equivalentes para todas las alternativas. Una de sus ventajas es que permite comparar los resultados independientemente de las unidades en las que se presenten. Sin embargo, cuando los proyectos presentan múltiples resultados, esto implica la necesidad de buscar una unidad común de comparación, como la hallada en el análisis coste-beneficio. Otra de las limitaciones de esta técnica, así como de la mayoría de técnicas cuantitativas, reside en la dificultad de ponderar los resultados que tienen una dimensión cualitativa importante (como sucede en el caso sanitario al valorar los resultados de salud en dimensiones como el dolor, por ejemplo).

[2] Note aquí el lector que no existe el análisis coste-eficiencia. Esto es debido a que el indicador de eficiencia ya contiene una relación entre los recursos y los resultados y por lo tanto ya sería un indicador completo en cuanto al proceso.

Sumario

El *Análisis Coste-consecuencia* es una variante del análisis coste-efectividad, y consiste en analizar el impacto de un nuevo programa en los costes y resultados esperados, presentando los diferentes componentes de forma desagregada y sin realizar juicios de valor acerca de la importancia relativa de cada uno de ellos. Esto permite al decisor hacer su propia apreciación de la importancia relativa de cada uno de los valores. Los componentes habituales de este análisis son: costes

Cuadro n.º 2

Métodos y técnicas de valoración monetaria

(A) Métodos de preferencia revelada
Obtención de valores monetarios implícitos en transacciones observadas en mercados reales donde alguno de los atributos del bien o servicio objeto de intercambio está relacionado con el atributo/bien/servicio que se quiere valorar
-Precios hedónicos
Estimación de la contribución marginal implícita del atributo relacionado con el nivel de satisfacción al precio del bien que es objeto de la transacción (salarios, viviendas, etc.). Referencia de Ej: Bilbao, C (2004)
-Coste del viaje
Estimación del valor monetario del coste que los individuos están dispuestos a soportar por acceder a un servicio en términos de tiempo y coste de desplazamiento. Referencia de Ej:Riera, A (2000, 2004)
-Costes evitados
Estimación de los costes necesarios para mitigar o reducir unos determinados efectos negativos sobre el bienestar social o individual. Referencia de Ej: Mogas, J (2004)
-Aportaciones voluntarias
Estimación basada en la disposición mostrada por los individuos de contribuir a organizaciones no lucrativas para finalidades relacionadas con la mejora del bien o servicio. Referencia de Ej: Menges R (2005)
(B) Métodos de preferencia declarada
Obtención de valores monetarios de la disponibilidad a pagar mediante escenarios o mercados hipotéticos.
-Valoración contingente
Estimación de la disponibilidad a pagar o la disposición a ser compensado (aceptar) mediante la simulación de un mercado hipotético con técnicas de encuesta. Referencia de Ej: Álvarez M et al (2003), Costa-Font J et al (2005)
-Análisis conjunto
Estimación de la disponibilidad a pagar a partir de la clasificación ordinal de diferentes alternativas mediante técnicas de encuesta. Referencia de Ej: Rivera A (2004)

Fuente: Adaptado de Puig-Junoy J. et al (2001)

Sumario

directos de la actividad, costes directos auxiliares, costes indirectos (coste de oportunidad y externalidades), impactos en resultados, e impactos en la intensidad e incertidumbre de los resultados.

El *Análisis Coste-beneficio* es quizás el instrumento más utilizado en todos los análisis, probablemente por su simplicidad teórica no exenta de complicaciones prácticas. Su ventaja principal es que permite realizar comparaciones entre proyectos similares o distintos, independientemente de las medidas de sus resultados, al homogeneizar la unidad de medida en valores monetarios. Este método podría definirse como el análisis en el cual todos los costes y beneficios se expresan en unidades monetarias, y es la ratio beneficios/costes o su diferencial en términos absolutos lo que determina si el proyecto es rentable o más/menos rentable que otro.

Sin embargo, no siempre es posible o deseable la transformación de los resultados (o de los costes no monetarios) en unidades monetarias. En algunos casos, no es posible debido a que no existe consenso científico acerca de su valor, o cuando este consenso no es política y/o socialmente aceptable (por ejemplo, en el valor monetario de una vida). En estos casos, la utilización del análisis coste-utilidad puede resultar de gran ayuda.

El *Análisis Coste-Utilidad* consiste en buscar una unidad de resultado que refleje el beneficio social de un programa público. Esta herramienta está especialmente extendida en el campo sanitario, en el cual existe suficiente consenso sobre su utilidad y donde el análisis coste-beneficio resulta de más difícil aplicación. El indicador de utilidad principal es el conocido como: «años de vida ajustados por calidad» (AVAC). Su principal utilidad reside en que es independiente de la renta de las personas, y permite comparar el resultado de dos tratamientos cuya efectividad puede ser muy distinta. Su aplicación más conocida es la que permite la creación de listas de prioridad de tratamientos sanitarios.

El *Meta-análisis* es una técnica de búsqueda sistemática, evaluación, combinación e integración de conclusiones de diferentes estudios independientes con el mismo efecto o resultado que la política que se analiza. El objetivo habitual consiste en obtener muestras más grandes cuando el coste impide hacerlo de forma rutinaria. Obviamente, este tipo de análisis está muy sujeto a la calidad de la información recogida, a la independencia de los estudios utilizados y al sesgo de publicación, más o menos presente (los estudios positivos tienden a públicarse más que los negativos). En este sentido, se trata de una técnica aplicable a todos los tipos de análisis, en función de los estudios de origen (coste-eficacia, coste-efectividad, coste-beneficio...).

Finalmente, a pesar de no haberlo incluido en ninguna de las categorías anteriores, es importante no olvidar el *análisis de sensibilidad* como parte final de cada uno de los análisis de proyectos públicos. Este consiste en ver cómo varían los resultados de la evaluación efectuada ante cambios producidos en los parámetros del análisis.

2.1.3. Técnicas aplicables a los análisis

Finalmente, un tipo de clasificación diferente responde a la aproximación empírica a la evaluación, según se utilicen téc-

nicas paramétricas, técnicas semi-paramétricas o técnicas no paramétricas.

A continuación analizaremos con más detalle algunas de las técnicas que se recogen más adelante en el cuadro n.º 3.

Técnicas paramétricas

Como su nombre indica, las técnicas paramétricas son aquellas que asumen que los datos están obtenidos a partir de una distribución funcional determinada.

Uno de los métodos paramétricos más utilizados para el análisis microeconómico de las políticas públicas es el de las fronteras de eficiencia (estocásticas y deterministas). Por eficiencia se entiende el grado de consecución de los objetivos al menor coste posible. Por lo tanto, se tiene en cuenta tanto el aprovechamiento de los recursos y su productividad, como el coste relativo de los recursos utilizados. Así pueden distinguirse dos conceptos de eficiencia, la eficiencia productiva —que queda definida por el concepto que acabamos de expresar—, y la eficiencia social. El concepto de eficiencia social nos permite extender el alcance de la eficiencia a escala social porque no es suficiente asignar correctamente los recursos y utilizar los de menor coste, sino que las decisiones deben tener la relevancia socialmente deseada. La eficiencia productiva puede dividirse entre Eficiencia técnica, que considera tan sólo la productividad de los factores, y Eficiencia económica o asignativa, que considera el coste relativo de los factores de producción.

Existen diferentes métodos para el cálculo de la eficiencia, y pueden distinguirse dos grandes grupos: aquellos basados en la estimación de una frontera de producción o de costes y que, por lo tanto, fijan la producción máxima o los costes mínimos posibles dado un entorno definido, y aquellos basados en el comportamiento medio. En cuanto a las técnicas desarrolladas, han evolucionado en torno a dos grupos diferentes: las técnicas paramétricas, que realizan una estimación de la función de producción y coste, y las no paramétricas, que calculan la eficiencia en función de las posiciones relativas de cada unidad respecto a las del resto del grupo (es decir, a las que se sitúan a la frontera, o a las que se sitúan en el comportamiento medio, en cada método, respectivamente).

Dentro de las técnicas paramétricas de cálculo de la eficiencia destaca la de fronteras estocásticas (Bateese, GE, Coelli TJ, 1992). En esta técnica, a partir de una forma funcional de costes o producción, se determina la distancia a la cual se encuentran el resto de unidades de producción. Esta distancia es considerada como la ineficiencia; en los métodos de frontera estocástica, la ineficiencia de la unidad de producción se separa de los efectos ambientales (efectos externos que varían a lo largo del tiempo) mediante la inclusión de un segundo término de error que varía entre unidades y a lo largo del tiempo. Las fronteras estocásticas requieren la disponibilidad de datos de panel (series temporales y transversales), a diferencia de las fronteras determinísticas (Aigner DJ, Chu SF, 1968) y de las técnicas no paramétricas como el DEA, que únicamente precisan datos transversales al incluir solo un término de error.

El análisis de la eficiencia asignativa también puede considerarse como una técnica paramétrica, al derivar los parámetros a partir de las fronteras de coste

(eficiencia global) y de los precios relativos, para compararlos con los obtenidos en las fronteras de producción (eficiencia técnica) y aislar así este último componente (eficiencia asignativa). Para ello se utilizan las «funciones de distancia» de Shephard (1953 y 1970).

Dentro de las técnicas paramétricas encontramos, también, el *análisis multivariante* utilizado para calibrar características que pueden ser distintas entre el grupo participante y el grupo control. Este método suele utilizarse en el caso de los diseños no experimentales.

Técnicas semiparamétricas

La mayoría de técnicas de evaluación requieren la disponibilidad de un conjunto de Unidades Técnicas de Producción (en adelante UTE) que permitan obtener indicadores a partir de la comparación entre ellas (técnicas no paramétricas como el DEA), o la obtención por estimación de una función de producción o costes con la cual comparar la unidad que se esté analizando. Sin embargo, estas condiciones no se dan en todas las situaciones. En la mayoría de programas aplicados a determinados sectores (salud y educación, entre otros) existe un gran interés por evaluar las experiencias piloto o pioneras para analizar su posible implantación en todo un territorio o en territorios distintos al analizado. Esto implica que, sin otro tipo de técnicas, dichas experiencias no podrían evaluarse más que en el espectro temporal o en relación con la situación previa a su introducción. Como veremos más adelante, cuando hablemos de los problemas más habituales de la evaluación, realizar una comparativa con la situación preexistente a la introducción de un programa puede suponer un grave error de medida, si existen (cosa muy probable) otros factores ambientales o internos a la propia organización que estén alterando (de forma positiva o negativa) los resultados obtenidos.

Las técnicas semiparamétricas como el *Matching* o las *Dobles diferencias combinadas con matching*. El *Matching* consiste en seleccionar factores suficientemente observables de forma que dos individuos con el mismo valor no presenten diferencias sistemáticas en reacción a la política pública. De este modo, si cada individuo puede aparejarse con un individuo que no ha participado en la reforma con las mismas variables de *Matching*, el impacto de dicha reforma puede analizarse a partir de la diferencia en sus efectos. El *Matching* puede, además, ponderar los resultados del programa según los condicionantes que han llevado a la unidad a participar en él, a través de la probabilidad (*propensity score*). Las *Dobles diferencias combinadas con matching* se utilizan para depurar los problemas que generan las dobles diferencias cuando se sospecha que, en alguna variable no observable, se está produciendo una selección de los individuos que participan en la reforma o política pública, respecto a los que no lo hacen.

El método de *variables instrumentales* consiste en utilizar variables que influyan en los motivos de la participación pero no en los resultados, de forma que se intenta controlar o aislar el efecto puro del programa o política. Este método se utiliza para datos que se han obtenido de forma no experimental y, de este modo, una vez corregidos los resultados según estas variables instrumentales, se observan los resultados corregidos.

Sumario

Técnicas no paramétricas

Dentro de esta categoría encontraríamos las técnicas de programación lineal[13] que no requieren la estimación de ningún parámetro, ya que simplemente comparan unidades de producción (o proyectos) comparables (actuales o pasados), y por lo tanto no determinan el «modelo ideal» sino que simplemente identifican los «mejores y peores resultados».

A continuación describiremos algunas de las características de estos mecanismos recogidos anteriormente en la tabla 1. Empezando por el Análisis de eficiencia no paramétrico con la técnica del Análisis Envolvente de Datos (DEA), que es un instrumento que permite la medición de la eficiencia técnica mediante la comparación de las unidades productivas de un sector (zona geográfica, departamento, etc.).

Quizás el mecanismo más sencillo, aunque de menor significatividad estadística, es el del *antes y después*, que consiste en comparar la media de las variables de resultado del grupo de individuos al cual se ha aplicado la política, antes y después de su aplicación. Este método requiere tener datos de antes y después de la reforma, pero sólo en el caso del grupo de participantes. Se suele aplicar cuando se tienen datos de panel, aunque también es posible aplicarlo cuando se dispone de un conjunto de secciones cruzadas o cortes transversales. Este método también puede ser paramétrico si se incorpora en un análisis de regresión que haga posible controlar las características de los individuos de la muestra.

La *Diferencia de medias (en panel o en sección cruzada)* consiste en obtener la media muestral de la variable resultado después de la reforma, para el grupo de tratamiento y el de control, y obtener la diferencia de estas medias. Una versión más refinada consiste en estimar una regresión en la que se incluye una variable indicador con valor 1, si el individuo está en el grupo tratamiento, y 0 en caso contrario.

Otra de las variantes de las técnicas no paramétricas es la de *Dobles diferencias o Diferencias en diferencias.* Se trata de una combinación de los dos métodos anteriores, y por tanto recoge ventajas e inconvenientes de ambos. Esta técnica está basada en la comparación de la media muestral de la variable resultado obtenida en el grupo de tratamiento y en el grupo control después de la aplicación del proyecto o la reforma. La diferencia entre ambas medias de antes y después del tratamiento nos dará tanto el sentido como la medida del resultado de la política.

Entre las técnicas no paramétricas también encontramos algunos métodos de cálculo de la eficiencia tales como el análisis envolvente de datos (AED o DEA) antes mencionado. Los AED constan de diferentes tipologías; según se orienten a los recursos (comparación de recursos utilizados dado un mismo nivel de producción) o a los resultados (comparación de los resultados obtenidos dado un mismo nivel de recursos); y según se utilicen las distancias radiales o no, es decir según se considere que la ineficiencia viene dada por un exceso proporcional de los recursos (radial), o aún siendo una uni-

[3] La programación lineal, una variedad de la programación matemática está compuesta por una función objetivo lineal, un conjunto de desigualdades lineales restrictivas y otro de condiciones de no negatividad.

dad ineficiente se pueda considerar que la cantidad utilizada de un recurso ya es óptima (no radial).

Se basan en el cálculo por programación lineal de las distancias de las unidades productivas a la frontera de mejor práctica. La frontera de mejor práctica está formada en este método por los valores de las unidades de producción con mejores resultados (orientación a resultados) o menores recursos utilizados (orientación a los recursos) y sus respectivas combinaciones lineales. Los Free Disposal Hull son una variante de los DEA, FDH (tipo de DEA) en el cual la frontera sólo está formada por las unidades reales (no sus combinaciones lineales), lo cual obliga a disponer de muchas más observaciones (Tulkens, 1993).

2.1.4. Mecanismos adicionales de control en la evaluación económica

La mayoría de técnicas de evaluación económica presentan resultados en función de algunas variables inciertas y datos procedentes de una muestra de la población, por lo que el resultado ofrece un cierto grado de incertidumbre. Para poder comprobar la significatividad de los resultados obtenidos, esta incertidumbre puede examinarse mediante algunas técnicas como el análisis de sensibilidad, o los tests estadísticos. En el caso de las pruebas estadísticas, sin embargo, si los datos de la muestra están sesgados (algo habitual cuando, por ejemplo, un grupo pequeño de usuarios puede generar unos hechos adversos muy costosos), o la muestra es muy pequeña, los tests habituales pueden conducir a resultados erróneos. En estos casos es recomendable la utilización de la técnica no paramétrica del *Bootstraping.* Esta técnica implica la simulación de observaciones de forma aleatoria, y de esta manera se evita el posible sesgo.

Otra observación importante que debe tenerse en cuenta es la presentación de resultados tanto en forma de ratio medio como de ratio incremental, cuando el proyecto se lleva a cabo agregándolo a otros proyectos preexistentes con objetivos similares o idénticos.

2.2. **Análisis macroeconómicos de evaluación de políticas**

En un segundo ámbito, bastante distinto a los anteriores, cabe reseñar las técnicas y tipos de análisis de políticas públicas que utilizan datos a nivel agregado y que buscan de forma global el impacto de las políticas sobre la economía.

El primero de estos análisis es el de las Tablas *Input-Output*. Su objetivo principal es evaluar el impacto que una política de un determinado nivel de gobierno (local, nacional, internacional) tiene sobre los agregados macroeconómicos del área (producción, empleo y renta). El diseño de estos estudios se basa en la interrelación de las unidades económicas a partir de una matriz de doble entrada, desagregada por ramas de actividad. Así, las columnas recogen los recursos o costes utilizados por cada rama de producción *(inputs)*, y las filas presentan su distribución entre consumos intermedios de otras ramas o consumos finales *(outputs)*. Esta distribución resulta en tres bloques o matrices: el de consumos intermedios o relaciones entre sectores, el de demanda final —consumos cuyo destino es la demanda final—, y el de recursos o in-

puts primarios, que muestra la aportación de cada rama o sector primario a la actividad o proceso productivo.

En segundo lugar podrían incluirse todas aquellas técnicas de evaluación de políticas públicas que se basan en la utilización de modelos de equilibrio general (en vez de equilibrio parcial, mucho más propio de las técnicas descritas como de análisis microeconómico en este texto), y de simulación. Dichas técnicas de evaluación se basan en el análisis de los agregados económicos y, en los casos de simulación, en la aplicación de los modelos de comportamiento teórico de la economía. De hecho, estos modelos utilizan las técnicas de la teoría macroeconómica para aplicarlas al análisis microeconómico de las políticas públicas. Las ventajas de estos modelos respecto a los de equilibrio parcial residen en que permiten captar los efectos de segundo orden o de realimentación.

Dentro de las técnicas de equilibrio general encontramos distintos modelos; para el análisis de políticas públicas destacan los Modelos de Equilibrio General Aplicado (MEGA), derivados de las teorías de Arrow y Debreu (1954). Otros tipos de análisis de equilibrio general son los modelos macroeconómicos y de Vectores Autorregresivos (VAR), que permiten analizar los efectos cruzados entre las variables objeto de estudio a lo largo de un periodo de análisis. Los modelos MEGA se diferencian de los anteriores en que utilizan datos de un sólo año de la Contabilidad Nacional. Se identifican como modelos de primera generación (utilizan el supuesto de competencia perfecta en un escenario estático), de segunda generación (también en un escenario estático consideran la existencia de rendimientos crecientes y competencia imperfecta), y de tercera generación (incorporan el concepto dinámico, ya sea de tipo Ramsey, con consumidor de vida infinita, o de generaciones solapadas). Los modelos de equilibrio general aplicado se utilizan para la simulación de políticas económicas tales como políticas fiscales, medioambientales, etc.

Finalmente, hay que destacar las posibilidades que ofrecen los nuevos modelos que intentan integrar ambas técnicas (Labandeira X, et al, 2006), o combinar técnicas distintas del mismo análisis macro (Labandeira, X., Labeaga, J., 1999), como, por ejemplo, las Tablas *input-output* con los modelos de equilibrio general.

2.3. La evaluación económica de políticas públicas en nuestro país

En nuestro país, las técnicas de evaluación económica de proyectos públicos han sido utilizadas por los analistas en todas las áreas de la actuación pública. Especialmente en los ámbitos propios del gasto social del estado del bienestar (educación, sanidad y servicios sociales), pero también en aquellas áreas más relacionadas con la gestión y el desarrollo de infraestructuras públicas.

Existen diferentes textos legales que exigen la realización de estudios que realicen análisis económico de proyectos públicos o de textos legislativos. Por ejemplo, la Ley 13/1989, de 14 diciembre de la Generalitat de Catalunya, de Organización, procedimiento y régimen jurídico, que en su artículo 63 incluía la necesidad de realizar análisis coste beneficio, y que se refunde tras la modificación por artículo único de Ley 4/2001, de 9 abril

(LCAT 2001, 234), de la siguiente forma: «La propuesta de disposición ha de ir acompañada de una memoria, la cual ha de expresar en primer lugar el marco normativo en el que la propuesta se inserta, ha de justificar su oportunidad y la adecuación de las medidas propuestas a los fines que se persiguen, ha de valorar la perspectiva de igualdad de género y ha de hacer referencia a las consultas que pueden haberse formulado y a otros datos de interés para conocer el proceso de elaboración de la norma. A la propuesta de disposición también se han de adjuntar: a) Un estudio económico en términos de coste-beneficio. [...]»; otro caso muy explicito es el de los Fondos Europeos (Reglamento Marco de los Fondos Estructurales) (Viñas V, 2002)

A continuación se exponen los resultados de una revisión sistemática de todos los artículos clasificados en la base de datos de economía (ECONLIT) durante los últimos 15 años, según los criterios que se detallan en el Anexo 1, donde la búsqueda se ha limitado principalmente a los modelos de tipo microeconómico. La información obtenida de dicha revisión se resume en la tabla siguiente, que a su vez sirve de síntesis de las diferentes técnicas y métodos descritas en el artículo. (Ver cuadro 3).

3. A MODO DE CONCLUSIÓN. LOS PROBLEMAS MÁS USUALES EN LA EVALUACIÓN ECONÓMICA, Y ALGUNAS RECOMENDACIONES

En esta última sección se analizan los motivos más habituales de distorsión de los resultados del análisis a partir de las experiencias reseñadas anteriormente, y se concluye con algunas guías o simples recomendaciones.

La no existencia de un elemento de comparación (contrafactual)

Tradicionalmente, la mayoría de estudios de evaluación económica que han evaluado proyectos de carácter singular han considerado que los efectos, en caso de no haberse desarrollado dichos proyectos, habrían sido nulos (cero costes, cero beneficios). Sin embargo, ello no es siempre cierto y, por tanto, en la mayoría de casos es necesario disponer de un contrafactual no nulo.

En la literatura existen diferentes aproximaciones acerca de cómo calcular este contrafactual, destacadas por Blundell y Costa (2004), y que podríamos resumir en a) experimentos sociales, b) experimentos naturales, c) métodos de matching, d) métodos instrumentales, y e) funciones de control.

Limitaciones de los mecanismos

Los diferentes métodos de evaluación económica no solo presentan problemas a la hora de realizar el ejercicio práctico, sino también algunas limitaciones en su aplicación. Algunas de estas dificultades ya han sido descritas a lo largo del artículo, como por ejemplo las implicaciones éticas de la valoración monetaria de la salud, o las dificultades para establecer pruebas piloto en las políticas públicas si se discrimina a una parte de la población sin tener criterios objetivos para realizar los experimentos.

A continuación se detallan algunas dificultades adicionales.

(a) Juicios de valor y singularidad de las elecciones. Este es una de las limita-

Cuadro n.º 3

Clasificación de las técnicas cuantitativas de análisis de políticas públicas y reformas sociales.

Análisis de coste (Comparativa entre proyectos)	Análisis de resultados (Comparativa entre proyectos)	Comparativa entre coste y resultados		Tipo de técnicas aplicadas en la realización de los análisis
Tipo de Análisis	Tipo de Análisis	Tipo de Análisis		Técnica
Modelos de frontera de costes y de eficiencia enfocados al input Análisis de minimización de costes (o simplemente análisis de costes) Análisis observacional de costes basado en el diseño del estudio Análisis de la eficiencia asignativa	Análisis de eficacia, Análisis de efectividad Análisis de eficiencia, Modelos de frontera de eficiencia enfocados al output(estocásticos, deterministas) Análisis de Beneficios	Análisis Coste-beneficio (9, 12) Análisis Coste-utilidad Análisis Coste-efectividad (61) Análisis Coste-consecuencia Análisis Coste-eficacia Análisis observacional de costes y resultados basado en el diseño del estudio	Técnicas paramétricas	Modelos de frontera de costes: fronteras estocásticas(23, 27, 32, 34, 50, 53, 63), fronteras deterministas Metanálisis (10) Matching Metanálisis Análisis observacional de resultados basado en el diseño del estudio Análisis Multivariante (52) Análisis de variables instrumentales (3, 5, 19) Funciones de Distancia (22)
			Técnicas semi-paramétricas	Matching Dobles diferencias combinadas con matching. Análisis de la eficiencia asignativa por los coeficientes de las estimaciones paramétricas Análisis de eficiencia técnica por envolvente de datos (AED o DEA) (8, 9, 17, 25, 28, 31, 34, 39, 41, 42, 43, 44, 45, 48, 51, 57, 58, 60) Índice de Malquist (26) Free Disposal Hull (29, 55, 56)
			Técnicas no paramétricas	Dobles diferencias Diferencia de medias en sección cruzada Antes y después (1, 4) Modelos de Equilibrio General Aplicado (2, 6, 14, 18, 35) Tablas *Input-Output* (11, 59) Análisis observacional (21)

Fuente: Elaboración propia

Nota: La numeración entre paréntesis corresponde a los artículos de la revisión realizada que se anexan al final del documento.

Sumario

ciones clásicas en los estudios de coste-efectividad, coste-eficacia o incluso coste-beneficio al presentar una alternativa A como mejor que la alternativa B

(b) En el Análisis coste-beneficio surge una distorsión habitual generada por el hecho de que la disponibilidad a pagar depende de la renta de los individuos. Cuando los individuos participantes en el proyecto tienen un nivel de renta muy distinto, los resultados pueden verse afectados por este factor.

(c) En el Análisis coste-efectividad y en las medidas de efectividad, aparecen dos problemas habituales: la existencia de múltiples efectos con unidades de efectividad diferentes, y la obtención de resultados de efectividad en unidades distintas para dos políticas en competencia. Al analizar los resultados de este tipo de estudios, es preciso descartar la existencia de resultados (positivos o negativos) secundarios respecto al objetivo principal de las políticas, que podrían distorsionar la elección entre las alternativas.

(d) Resumiendo las ventajas e inconvenientes de los modelos de frontera estocástica respecto a los de programación lineal (DEA y FDH), las primeras pueden ofrecer resultados erróneos si la forma funcional elegida dista mucho de la función de producción real -o de costes-, o si la distribución real de los datos es muy diferente de la asumida. Por otro lado, la elección de métodos de programación lineal ofrece resultados muy pobres cuando la disponibilidad de datos es limitada y, a su vez, imputa todas las variaciones de producción o coste a ineficiencia de las unidades de producción, cuando podrían deberse a efectos aleatorios externos.

(e) La técnica de análisis input-output sólo permite un análisis estático de los resultados, lo cual exige el supuesto de que tanto los consumos como los resultados se mantengan constantes en el tiempo (supuesto bastante restrictivo para la mayoría de políticas). Su otra limitación es que parte de una valoración muy básica de los efectos multiplicadores, lo que conduce a la obtención de unos resultados que deben ser analizados con cautela.

(f) Otro tipo de limitaciones atañe a la elección del método de análisis y al objetivo del mismo (análisis parcial de resultados, frente al análisis total). Mientras la mayoría de estudios económicos utilizan técnicas de análisis de la eficiencia, la mayoría de políticas tienen también un impacto distributivo –y de equidad- que no puede evaluarse si no se utilizan modelos de equilibrio general. Sin embargo, estos modelos de equilibrio general no incorporan fácilmente la heterogeneidad de los individuos o agentes económicos. El primero de los dilemas que se presentan en la elección del estudio reside en sacrificar, pues, información en el nivel de detalle (utilizando el análisis macro), o en sus efectos multiplicadores (utilizando el análisis micro). Los modelos de tipo microeconómico descritos ofrecen resultados basados en el impacto inmediato o a corto plazo, y son independientes de las teorías económicas. Por este motivo, son los que habitualmente utilizan los gobiernos para analizar el impacto de las políticas públicas, especialmente de los proyectos con un alcance limitado (de ámbito territorial o poblacional). Los modelos de tipo macroeconómico, especialmente los de simulación, son mayoritariamente aplicados por parte de los

Sumario

gobiernos para analizar el impacto de las políticas fiscales, o de políticas cuyo impacto se espera que sea más generalizado y a largo plazo (como los proyectos o legislaciones medioambientales).

En cualquier caso, aún conociendo las limitaciones de la evaluación económica para la toma de decisiones públicas, creemos que las alternativas (su no-utilización) pueden acabar siendo mucho peores. Por lo tanto, aunque elaborar sobre la evidencia resulte costoso e incierto, ¿alguien realmente sensato desea hacerlo sobre la ignorancia?

REFERENCIAS BIBLIOGRÁFICAS

AIGNER DJ, CHU SF (1968): *On estimating the Industry production function*. American Economic Review LVIII: 826-839.

ALBI E (2000): *Público y privado. Un acuerdo necesario*. Ed. Ariel. Barcelona, 2000.

ÁLVAREZ M, GONZÁLEZ M (2003): «Modelización semiparamétrica y validación teórica del método de valoración contingente: Aplicación de un algoritmo genético», *Hacienda Pública Española / Revista de Economía Pública*, 2003, v. 0, nº 164, págs. 29-47

ARROW, K. Y DEBRUE, G. (1954): *Existent of equilibrium for a competitive economy*. Econométrica. Vol. 22.

BATEESE, GE, COELLI TJ (1992): «Frontier Production Functions, technical efficiency and panel data: with application to paddy farmers in India». Journal of Productivity Analysis, 3: 153-169.

BILBAO, C (2004): «Determinación de la demanda de características de vivienda. Una aplicación para los principales municipios asturianos». *Regional and Sectoral Economic Studies*, July-Dec. 2004, v. 4, nº 2, págs. 109-28

BLUNDELL R AND COSTA DIAS M (2004): «Alternative Approaches to Evaluation in Empirical Microeconomics». *Institute for Fiscal Studies*. June 2004.

CHELIMSKY E. (1985): *Old Patterns and New Directions in Program Evaluation: Patterns and Directions*, Washington, D. C., ASPA.

COSTA-FONT J, ROVIRA J (2005): «Eliciting Preferences for Collectively Financed Health Programmes: The "Willingness to Assign"» *Approach, Applied Economics*, August 2005, v. 37, nº 14, págs. 1571-83

DESGAGNE A, CASTILLOUX A-M, ANGERS J-F, ET AL (1998): «The use of the bootstrap statistical method for the pharmacoeconomic cost analysis of skewed data». *Pharmacoeconomics* 1998; 13 (5 part 1): 487-97.

FARRELL MJ (1957): «The measurement of productive efficiency». *Journal of the Royal Statistics Society*, Series A, CXX (3):253-290.

LABANDEIRA, X., LABEAGA, J. (1999): «Combining input-output analysis and micro-simulation to assess the effects of carbon taxation on Spanish households». *Fiscal Studies* 20 (3), 303-318.

LABANDEIRA X, LABEAGA J M, RODRÍGUEZ M (2006): «A Macro and Microeconomic Integrated Approach to Assessing the Effects of Public Policies». Documento de Trabajo FEDEA 2006-02.

MAUSKOPF JA, PAUL JE, GRANT DM, ET AL. (1998): «The role of cost-consequence analysis in healthcare decision-making». *Pharmacoeconomics* 1998; 13 (3); 277-88.

MENGES R, SCHROEDER C, TRAUB S (2005): «Altruism, Warm Glow and the Willingness-to-Donate for Green Electricity: An Artefactual Field Experiment». *Environmental and Resource Economics*, August 2005, v. 31, nº 4, págs. 431-58

MOGAS AMOROS, JOAN (2004): «Métodos de preferencias reveladas y declaradas en la valoracion de impactos ambientales». *Ekonomiaz*, 3er Trimestre 2004, v. 0, nº 57, págs. 12-29

Sumario

POLAND, O (1974): «Program Evaluation and Administrative Theory». *Public Administration Review*; 34.

PUIG-JUNOY J, PINTO JL Y ORTÚN V (2001): «El análisis coste-beneficio en sanidad», *Atención Primaria*, 27 (6): 422-427, 2001.

RIERA A (2000): «Modelos de elección discreta y coste del viaje. Los espacios naturales protegidos en Mallorca». *Revista de Economía Aplicada*, Invierno 2000, v. 8, nº 24, págs. 181-201

RIERA A, FARRERAS V (2004): «El método del coste de viaje en la valoración de daños ambientales. Una aproximación para el País Vasco por el accidente del Prestige». *Ekonomiaz*, 3er Trimestre 2004, v. 0, nº 57, págs. 68-85

RIVERA A, (2004): «Aplicación del Análisis Conjunto en la priorización de una lista de espera quirúrgica». *Cuadernos Económicos de I.C.E.*, Junio 2004, v. 0, nº 67, págs. 93-106

SUBIRATS J (1994): *Análisis de Políticas Públicas y eficacia de la Administración*. Ministerio de las Administraciones Públicas. Serie de Administración General. Madrid.

TULKENS H (1993): «On FDH efficiency analysis : some methodological issues and applications to retail banking, courts and urban transit». *The Journal of Productivity Analysis*, 4: 183-210.

VIÑAS V (2002): «Descripción y análisis de la teoría sobre seguimiento y evaluación aplicable a las actuaciones cofinanciadas por Fondos Europeos». *Gestión y Análisis de Políticas Públicas* 25; 3-21, Septiembre-Diciembre 2002.

Anexo

Resumen de las publicaciones sobre Evaluación Económica en España, de especial atención para la elaboración, diseño o implementación de políticas públicas:

1. Una evaluación de las prestaciones sociales en la lucha contra la pobreza en Andalucía y España: Un análisis comparativo.; Gonzalez Rodriguez, M. Rosario; Basulto Santos, Jesus; Cuadernos Economicos de I.C.E., octubre, 2004, v. 0, nº 68, págs. 195-212

2. Evaluación de una reducción de las cuotas empresariales a la Seguridad Social, a nivel regional, a través de un Modelo de Equilibrio General Aplicado: El caso de Andalucía. Cardenete Flores, M. A.; Estudios de Economía Aplicada, abril, 2004, v. 22, nº 1, págs. 99-113

3. Evaluación microeconométrica de las políticas públicas de empleo: Aspectos metodológicos. Duran, Christian; Hacienda Pública Española /Revista de Economía Pública, 2004, v. 0, nº 170, págs. 107-33

4. El Impuesto sobre las Ventas Minoristas de Determinados Hidrocarburos: Una evaluación de sus efectos económicos.; Romero Jordán, Desiderio; Sanz Sanz, José Felix; Hacienda Pública Española /Revista de Economía Pública, 2003, v. 0, nº 164, págs. 49-73

5. Evaluación económica de políticas educativas: Una ilustración con la Ley General de la Educación de 1970.; Arrazola, Maria; De Hevia, José; Hacienda Pública Española /Revista de Economía Pública, 2003, v. 0, nº 164, págs. 111-27

6. La aritmética de una Renta Básica Parcial para España: Una evaluación con EspaSim.; Mercader Prats, Magda; Hacienda Pública Española / Revista de Economía Pública, Special Issue 2003, v. 0, nº 0, págs. 185-209

7. Información, evaluación y competencia al servicio de una gestión eficiente de los servicios públicos; González-Páramo, José Manuel; Onrubia Fernández, Jorge; Papeles de Economía Española, 2003, v. 0, nº 95, págs. 2-23 (Contextual)

8. Eficiencia global de la producción pública. Una primera aproximación a su evaluación.; Rueda López, Nuria; Papeles de Economía Española, 2003, v. 0, nº 95, págs. 94-112

9. Evaluación frontera de la eficiencia en costes: Aplicación a los municipios de Cataluña.; Giménez García, Víctor M.; Prior Jiménez, Diego; Papeles de Economía Española, 2003, v. 0, nº 95, págs. 113-24.

Sumario

10. Aspectos clave de la evaluación de la eficiencia productiva en la educación secundaria.; Mancebón Torrubia, Maria Jesús; Muñiz Pérez, Manuel Antonio; Papeles de Economía Española, 2003, v. 0, nº 95, págs. 162-87

11. Evaluación Económica Sectorial de la Inversión en Infraestructuras del Transporte: Aplicación al Vector de Inversiones 1990-1998 en España.; Tarancon Moran, Miguel A.; Regional and Sectoral Economic Studies, julio-diciembre, 2002, v. 2, nº 2, págs. 23-52

12. La agregación de costes y beneficios en la evaluación de proyectos intergeneracionales: El valor actual neto multigeneracional.; Padilla Rosa, Emilio; Pasqual Rocabert, Joan; Hacienda Pública Española / Revista de Economía Pública, 2002, v. 0, nº 163, págs. 9-34

13. La evaluación de la educación Secundaria Obligatoria. Marchesi, Alvaro; Papeles de Economía Española, 2000, v. 0, nº 86, págs. 150-64

14. El sistema de pensiones y las proyecciones de población: Una evaluación cuantitativa. Rojas, Juan A.; Cuadernos Economicos de I.C.E., 1998, v. 0, nº 64, págs. 171-88

15. Evaluación del Programa Prever durante los ejercicios de 1997 y 1998; Cabetas Hernandez, Mercedes; Economía Industrial, 1998, v. 0, nº 4, págs. 189-92

16. El debate de la estandarización metodológica en la evaluación de proyectos sanitarios: El caso de la tasa de descuento; Antoñanzas Villar, Fernando; Rovira Forns, Juan; Juarez Castello, Carmelo; Papeles de Economía Española, 1998, v. 0, nº 76, págs. 159-64 (metodológico)

17. La restricción de las ponderaciones en el Análisis Envolvente de Datos: Una formula para mejorar la evaluación de la eficiencia. Pedraja Chaparro, Francisco; Salinas Jimenez, Javier; Smith, Peter; Investigaciones Economicas, mayo, 1994, v. 18, nº 2, págs. 365-80

18. Convergencia, pacto social y política fiscal: Una evaluación macroeconómica. Andrés, Javier; Molinas, Cesar; Taguas, David; Revista de Economía Aplicada, otoño, 1993, v. 1, nº 2, págs. 5-29

19. Evaluación del régimen de financiación de las Comunidades Autónomas a través de un análisis del gasto y la renta. Garcia-Mila, Teresa; McGuire, Therese J.; Revista de Economía Aplicada, invierno, 1993, v. 1, nº 3, págs. 5-26

20. El objetivo ambiental en la evaluación pública de proyectos de inversión. Santamaria Aquilue, Rafael; Investigaciones Economicas, mayo, 1991, v. 15, nº 2, págs. 355-81 (metodológico)

21. Modelos estadísticos y evaluación: tres estudios en educación per Anna Cuxart Revista de Educación, 323, págs. 369-394, 2000

22. El efecto de la regulación sobre la eficiencia productiva: Una aplicación a las compañías ferroviarias europeas. Rodríguez Álvarez, Ana; Fernández Blanco, Victor; Baños Pino, José; Investigaciones Económicas, mayo, 2005, v. 29, nº 2, págs. 351-78

23. Infraestructuras y eficiencia técnica: Un análisis de técnicas frontera. Delgado Rodríguez, Maria Jesús; Álvarez Ayuso, Inmaculada; Revista de Economía Aplicada, otoño, 2004, v. 12, nº 35, págs. 65-82

24. Eficiencia y despilfarro en la atención sanitaria. A propósito del infarto de miocardio. (With English summary.); Meneu, Ricard; Peiro, Salvador; Cuadernos Económicos de I.C.E., junio 2004, v. 0, nº 67, págs. 9-25 (metodológico)

25. Especialización y eficiencia en el sistema hospitalario español. Rodríguez López, Fernando; Sánchez-Macias, José Ignacio; Cuadernos Económicos de I.C.E., junio, 2004, v. 0, nº 67, págs. 27-47 (PDF)

26. Eficiencia y productividad de la distribución farmacéutica en España (1993-2000). García-Ochoa, Fernando Gascón; González Fidalgo, Eduardo; Cuadernos Económicos de I.C.E., junio, 2004, v. 0, nº 67, págs. 227-42 (PDF)

27. Capital publico y eficiencia productiva: Evidencia para la UE-15. Delgado Rodriguez, M. Jesus; Alvarez Ayuso, Inmaculada; Hacienda Pública Española / Revista de Economía Pública, 2004, v. 0, nº 168, págs. 27-46

28. La eficiencia en las administraciones locales ante diferentes especificaciones del output. Balaguer Coll, Maria Teresa; Hacienda Pública Española / Revista de Economía Pública, 2004, v

Sumario

29. Un modelo FDH para la medida de la eficiencia en costes de los departamentos universitarios. Giménez García, Victor M.; Hacienda Pública Española / Revista de Economía Pública, 2004, v. 0, nº 168, págs. 69-92

30. Las prestaciones por desempleo en España: Eficiencia y equidad. Canto Sánchez, Olga; Toharia Cortes, Luis; Hacienda Pública Española / Revista de Economía Pública, Special Issue, 2003, v. 0, nº 0, págs. 125-56

31. Eficiencia global de la producción pública. Una primera aproximación a su evaluación. Rueda López, Nuria; Papeles de Economía Española, 2003, v. 0, nº 95, págs. 94-112

32. Eficiencia técnica y acciones estructurales en las comunidades autónomas. Una aproximación paramétrica. Bosch, Nuria; Espasa, Marta; Sorribas, Pilar; Papeles de Economia Espanola, 2003, v. 0, nº 95, págs. 149-59

33. Aspectos clave de la evaluación de la eficiencia productiva en la educación secundaria. Mancebón Torrubia, Maria Jesús; Muñiz Pérez, Manuel Antonio; Papeles de Economía Española, 2003, v. 0, nº 95, págs. 162-87 (metodológico)

34. La eficiencia de las organizaciones hospitalarias. Cabasés Hita, Juan M.; Martín Marten, José J.; López Del Amo González, M. Puerto; Papeles de Economía Española, 2003, v. 0, nº 95, págs. 195-212

35. Políticas de vivienda: Eficiencia y equidad. Lopez Garcia, Miguel-Angel; Papeles de Economía Española, 2003, v. 0, nº 95, págs. 226-41

36. Las pensiones contributivas en España: Eficacia y eficiencia. Dizy Menendez, Dolores; Papeles de Economía Española, 2003, v. 0, nº 95, págs. 260-69

37. Eficiencia y eficacia de la justicia. Pastor, Santos; Papeles de Economía Española, 2003, v. 0, nº 95, págs. 272-305

38. Análisis de la eficiencia de las instituciones encargadas de la seguridad ciudadana. Algunas enseñanzas procedentes del ámbito de la Economía de Delito. (With English summary.); Diez-Ticio Duce, Amor; Mancebón Torrubia, Maria Jesús; Papeles de Economía Española, 2003, v. 0, nº 95, págs. 306-19

39. La eficiencia en la administración de los tributos cedidos: Un análisis explicativo. Esteller More, Alejandro; Papeles de Economía Española, 2003, v. 0, nº 95, págs. 320-34

40. Influencia de la imagen corporativa en la eficiencia de los servicios públicos: Un análisis empírico. Zoghbi Manrique de Lara, Pablo; Armas Diaz, Alberto; CIRIEC-Espana, Revista de Economía Pública, Social y Cooperativa, noviembre, 2002, v. 0, nº 42, págs. 187-206

41. Eficiencia y cambio técnico en DEA con datos de panel. Aldaz, Natalia; Millan, Joaquin A.; Revista de Economia Aplicada, otoño, 2002, v. 10, nº 29, págs. 163-76 (NOTA)

42. Regional Productivity of Spanish Agriculture in a Panel DEA Framework; Aldaz, Natalia; Millan, Joaquin A.; Applied Economics Letters, February 2003, v. 10, nº 2, págs. 87-90

43. Una aproximación a la eficiencia en la gestión de la Agencia Estatal de la Administración Tributaria. Jimenez Aguilera, Juan de Dios; Barrilao Gonzalez, Pedro Enrique; Papeles de Economía Española, 2001, v. 0, nº 87, págs. 221-28

44. Análisis de la eficiencia productiva de las instituciones de educación superior. Martinez Cabrera, Marcelino; Papeles de Economía Española, 2000, v. 0, nº 86, págs. 179-91

45. Eficiencia en la inspección de Hacienda. González, Xose Manuel; Miles, Daniel; Revista de Economía Aplicada, otoño, 2000, v. 8, nº 24, págs. 203-19 (NOTA)

46. Efectos sobre la equidad y eficiencia de la política de peajes de autopistas. Riera, Pere; Najera, Tonatiuh; Papeles de Economía Española, 1999, v. 0, nº 82, págs. 166-77 (metodológico)

47. La contribución de la investigación operativa a la mejora de la eficiencia en el ámbito sanitario. Serra de la Figuera, Daniel; Ramalhinho Lourenco, Helena; Papeles de Economía Española, 1998, v. 0, nº 76, págs. 216-26 (metodológico)

48. La eficiencia en la Administración de Justicia. Las salas de lo contencioso de los Tribunales Superiores de Justicia. Pedraja, Francisco; Salinas, Javier; Revista de Economía Aplicada, otoño, 1995, v. 3, nº 8, págs. 163-95

Sumario

49. Costes y eficiencia en el transporte publico de viajeros. Rus Mendoza, Gines; Lopez del Pino, Francisco; Rodriguez Millán, Francisco; Revista de Economía Aplicada, invierno, 1995, v. 3, nº 9, págs. 91-104

50. Eficiencia técnica en los servicios de protección contra incendios.; Cuenca, Alain; Revista de Economía Aplicada, otoño, 1994, v. 2, nº 5, págs. 87-109

51. La restricción de las ponderaciones en el Análisis Envolvente de Datos: Una formula para mejorar la evaluación de la eficiencia. Pedraja Chaparro, Francisco; Salinas Jiménez, Javier; Smith, Peter; Investigaciones Económicas, mayo, 1994, v. 18, nº 2, págs. 365-80

52. Eficiencia y competitividad en los servicios públicos: Algunas consideraciones relativas a la asistencia sanitaria. Lopez Casasnovas, Guillem; Wagstaff, Adam; Moneda y Crédito, 1993, v. 0, nº 196, págs. 181-218.

53. Hospital Costs in Catalonia: A Stochastic Frontier Analysis; Wagstaff, Adam; Lopez, Guillem; Applied Economics Letters, July 1996, v. 3, nº 7, págs. 471-74

54. Indicadores de eficiencia en el sector hospitalario; Casasnovas, Guillem Lopez; Wagstaff, Adam R., 1992, Department of Economics and Business, Universitat Pompeu Fabra, Economics Working Papers.

55. Public Capital Effects on Spanish Regions Productivity: A Non-parametric Approach (1965-1998); Maria-Dolores, Ramón; Hacienda Pública Española /Revista de Economía Pública, 2004, v. 0, nº 171, págs. 57-74

56. Un modelo FDH para la medida de la eficiencia en costes de los departamentos universitarios. Giménez Garcia, Victor M.; Hacienda Pública Española /Revista de Economía Pública, 2004, v. 0, nº 168, págs. 69-92

57. Long- and Short-Run Non-parametric Cost Frontier Efficiency: An Application to Spanish Savings Banks ; Prior, Diego; Journal of Banking and Finance, April 2003, v. 27, nº 4, págs. 655-71

58. Technical Efficiency and Economies of Diversification in Health Care ; Prior, Diego; Sola, Magda; Health Care Management Science, 2000, v. 3, nº 4, págs. 299-307

59. Ajuste y proyección de las tablas input-output en condiciones de coherencia estructural mediante optimización matemática. Tarancón Moran, Miguel Ángel; Revista de Análisis Económico, diciembre, 2003, v. 18, nº 2, págs. 67-95

60. Assessing the Efficiency of Spanish Social Services Centres: An Empirical Application of Data Envelopment Analysis; Ayala, Luis; Pedraja-Chaparro, Francisco; Salinas-Jimenez, Javier; Empirical Economics Letters, July 2004, v. 3, nº 4, págs. 195-205

61. Costs and Effectiveness of a Syringe Distribution and Needle Exchange Program for HIV Prevention in a Regional Setting; Cabases, Juan M.; Sanchez, Eduardo; European Journal of Health Economics, September 2003, v. 4, nº 3, págs. 203-08

62. Partitioning Input Cost Efficiency into its Allocative and Technical Components: An Empirical DEA Application to Hospitals, Socio-Economic Planning Sciences, V4-2/3, 1-20, 2000.

63. Cost Efficiency in Primary Care Contracting: A Stochastic Frontier Cost Function Approach. Puig-Junoy J, Ortún V, Documento de Trabajo Universidad Pompeu Fabra, 2001.

64. La eficiencia asignativa de RENFE durante el periodo 1955-1995. Baños-Pino et al. (2002)

65. Determinantes de la eficiencia técnica en el sector ferroviario europeo. Cantos et al. (1999)

Sumario

Métodos cuantitativos y «benchmarking»: su utilidad para orientar las políticas públicas

Un área en la que los métodos cuantitativos, entendidos en sentido amplio, han demostrado ser útiles por su contribución al diseño y evaluación de las políticas públicas es la elaboración de «ligas» o ranking, bien sean de criterio único o múltiple. Investigadores académicos y responsables de las políticas públicas están alcanzando puntos de encuentro en su aplicación a países, a unidades de producción y en la priorización de problemas. El artículo ofrece una tipología de los métodos y revisa su aplicabilidad, limitaciones y retos para el futuro.

Metodo kuantitatiboak, zentzu zabalean hartuta, baliagarriak izan dira arlo batean, politika publikoak diseinatzeko eta ebaluatzeko egin duten ekarpena dela-eta, hain zuzen, irizpide bakarreko edo askotako «ligak» edo rankingak egiteko arloan. Ikertzaile akademikoak eta politika publikoen arduradunak elkarguneak aurkitzen ari dira horiek herrialdeei eta ekoizpen-unitateei aplikatzeari dagokionez, arazoen lehentasunak ezartze aldera. Artikuluak metodoen tipologia eskaintzen du, eta haiek aplikatzeko modua eta haien mugak eta etorkizunerako erronkak berrikusten ditu.

The elaboration of ranking or league tables is a major area in which quantitative methods contribute to public policies design and evaluation. These rankings may be based on one or on multiple criteria. Academic researchers and those responsible for public policies are collaborating in the application of these methods to countries, departments and to problems priorization. This paper classifies the methods and reviews their applicability, their limitations and the challenges for the future.

Beatriz González López-Valcárcel
Universidad de Las Palmas de GC

ÍNDICE

1. Introducción.El contexto, los métodos cuantitativos y las políticas públicas
2. El paradigma del *benchmarking*
3. Ranking unidimensional: de la calidad de los proveedores a la comparación de «ciudades gemelas» pasando por la liga AVAC
4. Ranking multidimensional ¿Cómo ordenar cuando las políticas tienen múltiples objetivos?

Referencias bibliográficas

Palabras clave: evaluación, métodos cuantitativos, política pública

N.º de clasificación JEL: C10, I18, H5

1. INTRODUCCIÓN. EL CONTEXTO, LOS MÉTODOS CUANTITATIVOS Y LAS POLÍTICAS PÚBLICAS

El diseño y la evaluación de las políticas públicas es objeto de una intensa atención científica, particularmente en los tiempos que corren, en los que el concepto de coste de oportunidad ha trascendido más allá de la academia y las sociedades son conscientes de la limitación de los recursos públicos y de la necesidad de elegir. Una política pública viene definida por uno o por varios programas, simultáneos o secuenciales, con objetivos que pueden ser complementarios o no. Está surgiendo un cuerpo de literatura científica dedicada a evaluar las políticas, con estrategias e instrumentos microeconométricos de identificación y estimación (Frölich, 2004). La problemática estadística que suscitan tales problemas entronca directamente con la posibilidad de experimentación (si se asignan los participantes a los programas aleatoriamente, sobre una población homogénea, la inferencia sobre los efectos medios de cada programa es inmediata). Dado que en ciencias sociales en general, y en políticas públicas en particular, no se puede experimentar tanto como los estadísticos desearían, por problemas políticos —asignaciones condicionadas por elementos contextuales—, por no perder eficiencia —conviene dar la beca al candidato con mejores expectativas de rendimiento— o por restricciones éticas —para probar la efectividad de un tipo de cirugía no es ético crear un grupo de control formado por pacientes que se so-

meten a falsa cirugía o cirugía placebo—, los métodos cuantitativos se han empeñado en identificar y estimar los resultados de las políticas públicas mediante diseños semi—experimentales, y mediante el uso de información exógena en estudios observacionales.

Se aplican métodos cuantitativos a todo tipo de políticas públicas. En este artículo, presentamos ilustraciones y casos provenientes mayoritariamente del ámbito de las políticas de salud, aunque hacemos referencia a otros ámbitos a los que la metodología puede extenderse de forma inmediata y sencilla.

Entendemos los métodos cuantitativos en sentido amplio, incluyendo los métodos estadísticos, econométricos, los algoritmos de computación, y la investigación operativa. Todos ellos aportan instrumentos válidos para diseñar y evaluar políticas públicas. En su aplicación al ámbito sanitario, se benefician de la simbiosis entre la economía y otras disciplinas (ciencia política, epidemiología,…), integrados en la Investigación en Servicios Sanitarios (ISS), la cual está tomando cuerpo y reconocimiento oficial como ciencia híbrida sin filiación disciplinar, ocupada de aunar perspectivas múltiples, académicas y profesionales, en un esfuerzo por mejorar los servicios sanitarios. Un proceso similar tuvo lugar después de la II Guerra Mundial con la Investigación Operativa.

Dentro de los métodos cuantitativos, la microeconometría ha experimentado un gran desarrollo en los últimos años, aportando a la econometría general y recibiendo de ella, con aportaciones específicas para el ámbito de la salud (Jones, 2000), de la economía laboral, de la educación, y de otras políticas públicas. El interés de los investigadores, movido por la necesidad de resolver cuestiones pero también por la oportunidad de obtención de datos y por la capacidad de cálculo de los ordenadores, se ha desplazado desde los modelos para datos transversales o para series temporales hacia los modelos para datos longitudinales (de panel) y los modelos para datos jerárquicos. Ahora que se dispone de registros poblacionales y grandes bases de datos, se trabaja con poblaciones completas, y los elementos de error aleatorio de los modelos reflejan errores de medida de las variables y errores de especificación de los modelos (la heterogeneidad no observable), más que errores de muestreo.

El uso generalizado de modelos para datos longitudinales con muestras grandes —el panel de hogares, por ejemplo— y de modelos multinivel con datos jerárquicos implica que se trabaja con varias dimensiones, cada una con su correspondiente tamaño muestral. Generalmente, los paneles tienen tamaños grandes en el espacio (n grande) pero seguimiento temporal corto (t pequeño). Paradójicamente, con los grandes paneles de microdatos seguimos sufriendo el problema de muestras pequeñas, aunque ahora lo son en la dimensión temporal (seguimiento corto de la población), por lo cual hay que buscar métodos de estimación consistentes en la dimensión temporal.

Se plantea un problema similar con los modelos multinivel, en los que hay tantos tamaños muestrales como niveles. Por ejemplo, si el nivel 1 son los pacientes y el nivel 2 los hospitales en donde ingresan, los tamaños muestrales relevantes son el número de pacientes ingresados en cada hospital y el número de hospita-

Sumario

les. Lo propio ocurre en los modelos que evalúan el rendimiento educativo bajo diferentes políticas o programas. El nivel 1 es el alumno, el nivel 2 es el centro donde estudia. De hecho, los modelos multinivel nacieron en el ámbito de la educación, para evaluar comparativamente el efecto de las capacidades individuales y de la influencia de la escuela en el éxito educativo (Goldstein, 2003).

La investigación biomédica y en servicios sanitarios tiende a generar datos experimentales y a aplicar una metodología estadística que ya está bastante estandarizada (casos-controles, estudios de cohortes). Hay protocolos metodológicos para los ensayos clínicos (criterios de selección, diseño estadístico del experimento) que están bien desarrollados. En econometría seguimos siendo generalmente usuarios pasivos de datos no experimentales. No pudiendo experimentar con la población, se experimenta con la muestra, por ejemplo mediante los métodos bayesianos de estimación (*Gibbs sampling, Bootstrapping, MCMC*).

Por otra parte, hay desarrollos recientes de métodos microeconométricos para evaluar los resultados de programas públicos, que fundamentalmente se están aplicando en economía laboral (Angrist et al., 1998;Angrist et al., 2001), aunque tienen un gran campo abierto de aplicación en economía de la salud (Vera-Hernandez, 2003). Tratan de estimar el efecto medio de cada tratamiento (resultado esperado en caso de participación en el programa menos resultado esperado en caso de no participación; por ejemplo, el resultado es el salario y el tratamiento un programa formativo de fomento del empleo) con modelos paramétricos o no paramétricos que resuelvan, entre otros, los problemas de sesgo de selección (la adscripción al programa no es aleatoria), y descuentan el efecto de los factores de confusión. En la actualidad se están desarrollando modelos microeconométricos para evaluar programas que emplean múltiples tratamientos (o intervenciones políticas simultáneas) (Frölich, 2004). Esperamos que en los próximos años la economía laboral y la economía de la salud converjan y se polinicen mutuamente, al menos en lo que atañe a ese tipo de modelos.

La estandarización metodológica de los métodos cuantitativos en el ámbito de las políticas públicas responde a un nuevo paradigma, el de la Política Basada en la Evidencia (PBE), derivado por contagio de la Medicina Basada en la Evidencia (MBE). Se practica el *benchmarking* para aprender de los demás, comparar y estandarizar reformas y para asignar recursos centralizados a las unidades prestadoras de los servicios asistenciales.

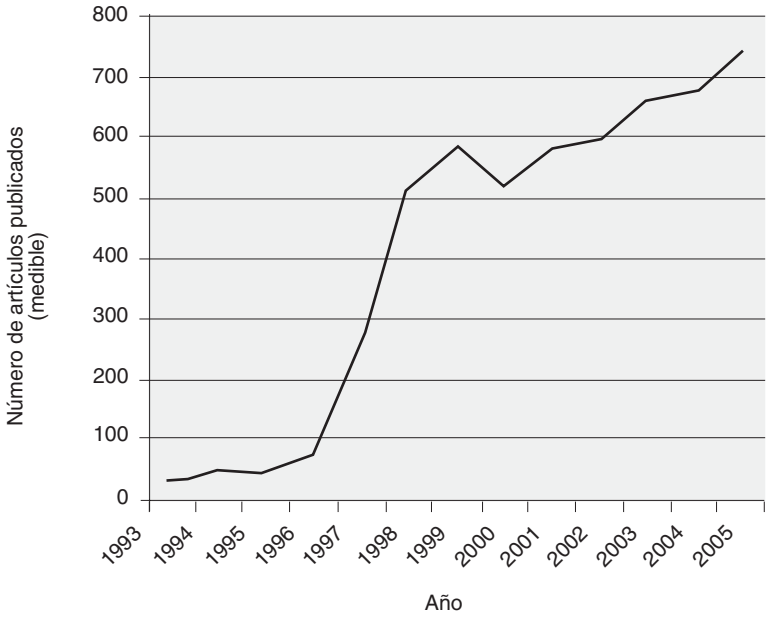
En este artículo presentamos una visión general del uso de métodos cuantitativos para construir *ranking* que orienten las políticas públicas, con especial referencia a las de salud, y cómo emplearlos. Se hace hincapié en los problemas metodológicos y en las dificultades para mejorar las políticas a partir de los análisis, más que en lo intrincado de la metodología estadística o en los resultados concretos de los ejemplos que se presentan.

2. EL PARADIGMA DEL *BENCHMARKING*

Vivimos en un mundo obsesionado por la e*videncia* que busca hacer Políticas Basadas en la Evidencia (PBE). El *bench-*

Figura nº 1

El *benchmarking* en las publicaciones médicas 1993-2005



marking se practica tanto a escala de mesogestión de centros- Top20 hospitales(Garcia-Eroles et al., 2001;Iasist, 2006)- como a escala mundial. La Organización Mundial de la Salud (OMS) compara los sistemas sanitarios en cuanto a los recursos que emplean y a los resultados que consiguen. La OCDE ha puesto en marcha sistemas de indicadores, y bases de datos internacionales, para los ámbitos sanitario, educativo y otros. Las Naciones Unidas elaboran y difunden, ya desde hace años, el Indice de Desarrollo Humano que combina varias dimensiones del bienestar —PIB, educación, salud— Gracias a los indicadores homogéneos internacionalmente de la OCDE (Kelly et al., 2006; OECD, 2003) podremos, por ejemplo, comparar el «tiempo puerta-aguja» en caso de infarto en España con el de los vecinos, y orientarnos por *benchmarking* para encontrar el camino de las intervenciones eficientes. El *benchmarking* se usa también ampliamente para las siguientes actividades: diseñar sistemas de incentivos a los gestores; ajustar fórmulas de pago a proveedores (separar el grano de la paja, no remunerar la parte de los costes que corresponde a la ineficiencia); elaborar ligas de hospitales o de universidades que informen al público y permitan elegir con criterios de calidad; ordenar posibles intervenciones públicas según criterios de coste-efectividad (ligas de AVACs), entre otras aplicaciones.

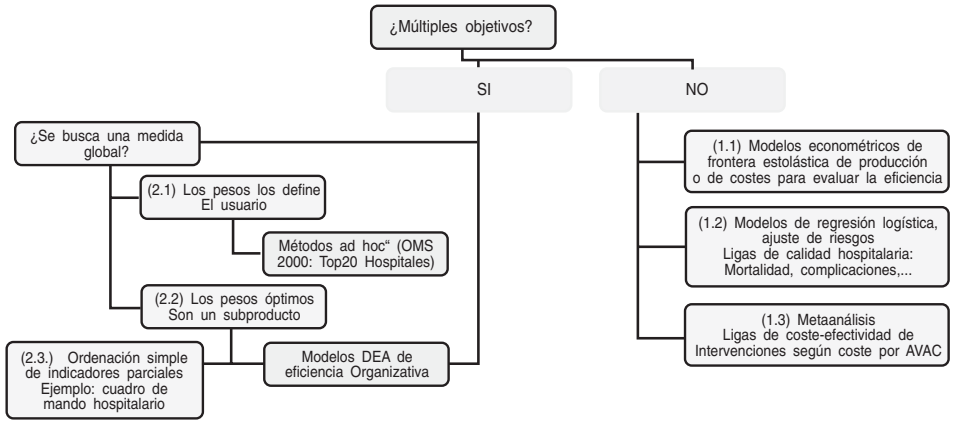
Sumario

El *benchmarking* como instrumento de mejora sistemática de las organizaciones procede de la literatura de la Gestión de Calidad Total. Empezó a extenderse hacia las publicaciones científicas del ámbito sanitario a mediados de los noventa, aunque no se incluye como término de referencia en *medline* (*Medical Subject Headings*) hasta 1998, con la siguiente definición: «método para medir el desempeño (*performance*) contra estándares establecidos de la mejor práctica». En ese período se iniciaron algunos experimentos puntuales (Nelson et al., 1995). La progresión de artículos publicados en los últimos diez años en el ámbito sanitario (figura 1) ha sido notable. El gran salto se ha producido entre 1996 y 1998.

Los métodos para obtener un *ranking* difieren según los objetivos, los datos y el modelo subyacente (grado de incertidumbre y forma de incorporar, en su caso, juicios de valor para ponderar diferentes objetivos). Desde el punto de vista estadístico es relevante la diferenciación entre modelos deterministas y aleatorios, pero desde la perspectiva de las políticas es más interesante la clasificación de la figura 2. En ella diferenciamos los *ranking* que corresponden a uno y a múltiples objetivos, y entre éstos separamos los métodos que ponderan exógenamente los objetivos, a criterio del modelizador, y los que obtienen pesos óptimos de cada uno de los objetivos —u *outputs*— como resultado, y no como *input*, del método.

Figura nº 2

Los ranking en sanidad. Una clasificación de los métodos orientada a las políticas



Fuente: Elaboración propia

Sumario

3. RANKING UNIDIMENSIONAL: DE LA CALIDAD DE LOS PROVEEDORES A LA COMPARACIÓN DE «CIUDADES GEMELAS» PASANDO POR LA LIGA AVAC

Los *ranking* de criterio único incluyen tres tipos de teorías bien diferenciadas: 1) los modelos econométricos de frontera estocástica de producción o de costes, que tratan de medir comparativamente la eficiencia de las unidades productivas, en el marco teórico de la teoría económica de la producción; 2) los modelos de regresión logística y similares, que estiman la calidad (por ejemplo, mortalidad hospitalaria, implicaciones, reingresos,… esperada de cada centro, ajustando por severidad de los casos tratados, y en su caso por otros factores no controlables de entorno); 3) los *ranking* de intervenciones según coste-efectividad, que tratan de comparar el coste de oportunidad de programas alternativos, para ayudar a los financiadores públicos en la priorización informada de los problemas y de las intervenciones sobre ellos.

Cada vez con mayor frecuencia y trascendencia financiera se elaboran, publican y utilizan ordenaciones de unidades proveedoras de servicios públicos, por ejemplo hospitales, o centros de atención primaria, basados en un único criterio u objetivo.

Los modelos econométricos de frontera estocástica estiman el grado de ineficiencia productiva de un conjunto de unidades de producción. Estiman la ineficiencia técnica y/o asignativa, con fronteras de producción o de costes, suponiendo que la propia frontera es estocástica, por lo que son modelos con un «error compuesto» de dos componentes. Uno de ellos forma parte de la frontera y representa los errores de medida, omisión de variables y presencia de acontecimientos no predecibles y fuera de control que afectan a la producción. El segundo componente, asimétrico, (positivo en modelos de fronteras de costes, y negativo en modelos de fronteras de producción), capta la ineficiencia.

Un modelo de frontera estocástica de producción para un único *output* (y) de la unidad de producción i-ésima se formularía así:

$$\log (\mathbf{Y}_i) = x'_i\beta + z'_i y + v_i + u_i$$

Donde x´es un vector de k factores de producción (o una función de ellos, por ejemplo, el logaritmo), z´es un vector de variables de entorno no controlables que influyen en la producción, v es un error aleatorio que se supone repartido con una función de distribución continua de media cero independiente de las x y de las z, y u es la ineficiencia, variable aleatoria negativa o nula, distribuída independientemente de las x y de las z. El modelo de frontera de costes se formularía de forma similar, salvo que ui≥0. Estos modelos admiten diversas formas funcionales (González-López-Valcárcel et al., 1996). Aunque se han sofisticado en varias direcciones, no están todavía listos para ser empleados con propósitos prácticos de medir la eficiencia con fines remunerarativos. Su atractivo, hasta ahora, es más académico que profesional (Chirikos et al., 2000; González-López-Valcárcel et al., 1996;Jacobs, 2001;Worthington, 2006). En su aplicación a sanidad, una parte del problema es que están desvinculados de los resultados clínicos y de la calidad asistencial. Los ajustes por calidad son un problema común en la eva-

Sumario

luación de servicios públicos de múltiples *outputs*, donde el mercado no emite señales porque no se comercializan (las universidades o la policía serían equiparables a los hospitales).

Ranking unidimensional de la calidad (clínica) hospitalaria

Los *ranking* unidimensionales de calidad clínica de los hospitales, por el contrario, se emplean para emitir señales de calidad y orientar la elección de los pacientes, en un contexto de política sanitaria de autorregulación por el mercado. En Estados Unidos, ya desde hace años se elaboran y difunden *ranking* de calidad de hospitales para determinadas intervenciones, es el caso del tratamiento del infarto y de la cirugía cardíaca, incluso se comparan los resultados de cirujanos individuales (Green et al., 1995). En Nueva York se publicó el *ranking* de cirugía cardiaca por primera vez en 1997. Algunos de los servicios que quedaron en los últimos puestos, cerraron, el resto mejoró su tasa de mortalidad en los años sucesivos (Cutler et al., 2004). Se demostró que el factor experiencia influye en la mortalidad (Wu et al., 2004).

Desde la perspectiva de las políticas, son muy interesantes los experimentos en curso de pago por resultados (*pay-for-performance*, P4P) a los hospitales, que se decidió en EEUU en 2003 y se inició experimentalmente en 2005, para incentivar financieramente a los hospitales, dentro de los programas de aseguramiento público Medicare y Medicaid, por proveer atención sanitaria de alta calidad a sus pacientes ingresados. La base de datos, mantenida por los *Centers for Medicare and Medicaid Services* (CMS), contiene datos para unos 4.200 hospitales de 17 indicadores de calidad clínica del tratamiento del infarto agudo de miocardio, fallo cardíaco, neumonía y prevención de infección quirúrgica. La base de datos es pública y de libre acceso, permitiendo al usuario comparar la calidad de los hospitales de su zona (US Dept.Health and Human Services HHS, 2006). El experimento financiero, de momento, se ha limitado a unos 270 hospitales con participación voluntaria y los primeros resultados son alentadores (Kahn III et al., 2006).

La elección del criterio de calidad (mortalidad, complicaciones, reingresos) no es trivial. El indicador elegido debería ser fiable, válido, sensible, preciso, con interpretación clínica, útil tanto para la elección informada de hospital por los pacientes como para que los profesionales dispongan de estándares de buena práctica. Así mismo debería ser fácil de obtener y difícil de manipular.

Las ligas de calidad clínica de hospitales tienen detractores y entusiastas y han sido discutidas por dos tipos de problemas. En primer lugar, los asociados a la propia medida de calidad y su ajuste por factores de confusión, es decir, problemas de posibles sesgos sistemáticos relacionados con el modelo estadístico o econométrico que generó el *ranking* y los ajustes por riesgo. En segundo lugar, por problemas de significación estadística de las diferencias en las posiciones dentro del *ranking*. Puesto que los indicadores medidos son al fin y al cabo extracciones aleatorias de una distribución de probabilidad, hay errores de muestreo que además varían entre unidades. Puede ocurrir que la incertidumbre implícita en los datos sea tan elevada que de hecho no

Sumario

haya diferencias significativas entre centros que ocupan distintas posiciones en el *ranking*. Aunque hay contrastes diseñados para comparaciones múltiples (Bonferroni), la mayor parte de los *ranking* publicados no informan al respecto.

Veamos un ejemplo ilustrativo sencillo. Sean dos hospitales, uno es un pequeño hospital con 40 intervenciones anuales de cirugía cardíaca y el otro es un complejo hospitalario que hace 1.000 intervenciones al año. Para estimar la tasa de complicaciones (o de mortalidad, o reingresos) se calcula la frecuencia muestral (es decir, la proporción de pacientes intervenidos que sufrieron el suceso adverso; se suponen pacientes homogéneos, que no hay sesgo de selección y que ambos hospitales tienen una probabilidad idéntica de que ocurra el suceso adverso (igual calidad) y los números de sucesos adversos siguen sendas distribuciones binomiales con parámetros (n,p) donde n es 40 para el hospital pequeño y 1.000 para el grande, y p es idéntica.

Dados n y p podemos calcular los intervalos simétricos respecto a p que contienen una proporción determinada (1-α), por ejemplo el 95%, de la masa de probabilidad de la distribución respectiva. Los límites inferior y superior de las tasas del efecto adverso, p_1 y p_2, se determinarán para cada hospital mediante las expresiones siguientes:

$$\Pr(p_1 \le p \le p_2) = 1-\alpha$$

$$\sum_{x=0}^{np_1}\binom{n}{x}p^x(1-p)^{n-x} = \frac{\alpha}{2}$$

$$\sum_{x=np_2}^{n}\binom{n}{x}p^x(1-p)^{n-x} = \frac{\alpha}{2}$$

Hemos representado en la figura 3 dichos límites (p_1 y p_2) para ambos hospitales, para una probabilidad del 95% y niveles de calidad (tasas de efectos adversos) entre el 5% y el 25%. Como puede observarse, los intervalos son mucho más amplios para el pequeño hospital que para el grande. Además, las distancias se acrecientan para sucesos adversos más probables (al aumentar p).

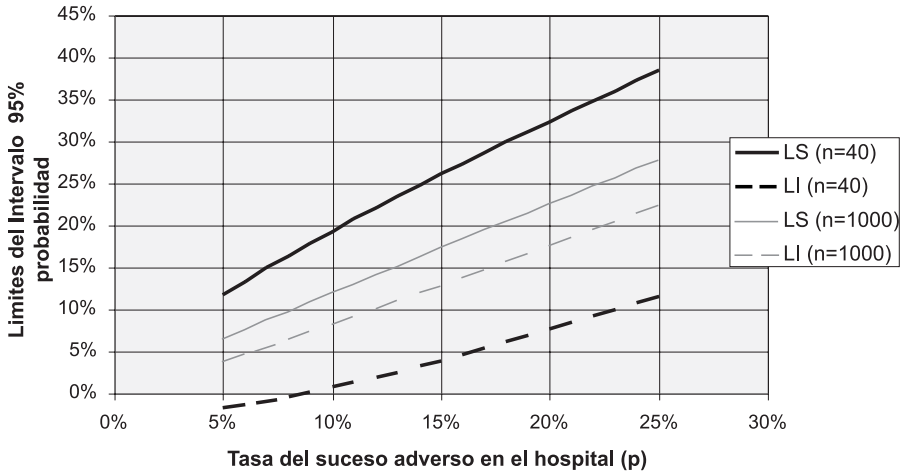
Si suponemos que el número de pacientes intervenidos (n_i) en cada hospital (i) es un parámetro fijo, el número de fallecidos en cirugía cardiaca es Binomial B(n_i,p_i). El parámetro de interés es p_i (tasa de fracasos, por mil intervenciones). Su IC dependerá del número de casos (n_i) y de la probabilidad de éxito y de fracaso (p_i y 1- p_i). Cuanto menor sea el tamaño o nivel de actividad de un hospital, menor será la potencia de los contrastes sobre su calidad. Los hospitales pequeños tienen intervalos de confianza amplios, compatibles con un amplio rango de calidad asistencial. En la práctica, pues, tendrán mayor probabilidad de ocupar los puestos en ambos extremos del *ranking*. Si las listas no informan sobre intervalos de confianza, están contaminadas por el efecto tamaño y las comparaciones pueden no ser relevantes.

Pero no sólo hay contaminación por el efecto tamaño. Es más grave el sesgo de selección y la heterogeneidad no observable entre pacientes (factores de confusión). Por ejemplo, la mortalidad de los pacientes ingresados con infarto de miocardio puede depender más del estado de salud en que llegan al hospital que de la calidad de la atención sanitaria que éste les presta. Puesto que hay sesgo de selección (los mejores hospitales reciben a los pacientes más graves), si no se ajusta por gravedad se encuentra mayor mortalidad asociada a los mejores hospitales.

Sumario

Figura nº 3

Efecto del tamaño muestral y del nivel de calidad (p) sobre la amplitud del intervalo que contiene el 95% de probabilidad



Nota: los límites se han calculado aproximando la distribución binomial a la normal

Para ajustar por dichos factores se emplean modelos de regresión, generalmente logística, que predicen la mortalidad de diferentes servicios hospitalarios homogéneos (por ejemplo, cirugía cardiaca) ajustando por gravedad y por otras variables de confusión, y ordenan a los hospitales en orden creciente o decreciente de calidad. Una solución elegante al sesgo de selección —hospitales más prestigiosos reciben pacientes más graves— proviene de los modelos econométricos bayesianos, que emplean la información exógena sobre distancia del domicilio del paciente a los distintos hospitales para estimar el sesgo de selección (Geweke et al., 2001). Mediante métodos de estimación bayesiana, por simulación (MCMC) Geweke y colaboradores obtienen distribuciones a posteriori de calidad y consiguen comparar la calidad de 114 hospitales americanos en el tratamiento de la neumonía: comparaciones bilaterales y entre grupos de hospitales.

En las ordenaciones subyace un proceso anidado, con al menos dos niveles: pacientes y hospitales. Los modelos multinivel son adecuados para este fin porque separan los efectos del grupo (hospital) de los efectos individuales (Goldstein et al., 2003). Ya se ha mencionado que su generalización a otros ámbitos de las políticas públicas es casi inmediata (alumnos y colegios, candidatos y programas, o procesados y jueces, por ejemplo).

Sumario

Desde la perspectiva econométrica, los modelos paramétricos de frontera estocástica y los modelos multinivel tienen en común que, frente a los econométricos tradicionales, consideran que la parte aleatoria del modelo (el «ruído») es información, y de hecho lo que interesa a efectos de políticas no es tanto el valor estimado de los parámetros fijos, coeficientes de las x, que son meros ajustes, sino las varianzas de los componentes del error, en los modelos de regresión multinivel (Goldstein, 2003;Snijders et al., 1999), o estimar (predecir) a partir de los residuos cada uno de los valores de los errores aleatorios de las unidades muestrales (ineficiencia). Este interés en estimar consistentemente los errores aleatorios a partir de los residuos se ha beneficiado de la metodología bayesiana.

En cualquier caso, debe imponerse la prudencia al leer los resultados de las «ligas», que a veces se presentan de forma excesivamente simplista, haciendo un uso inadecuado de la estadística (Vass, 2001). Por ejemplo, sin informar de los intervalos de confianza, o sin el necesario ajuste por riesgos. La publicación de estos datos contribuye a mejorar la calidad de los servicios cuando su nivel de calidad es bajo, como ha demostrado un estudio experimental sobre la seguridad de los hospitales de Wisconsin (Hibbard et al., 2003).

Los modelos de regresión logística que ajustan por riesgos una característica dicotómica de calidad con objeto de comparar hospitales se emplean también para comparar zonas geográficas. Puesto que la variabilidad intra país es mayor que la variabilidad entre países, se han comparado ciudades «gemelas» de distintos países (París y Manhattan, por ejemplo). Comparando las tasas de altas hospitalarias potencialmente evitables por la atención primaria(Brown et al., 2001;Sanderson et al., 2000) mediante sendos modelos de regresión logística para Paris y Manhattan, se ha determinado que el riesgo relativo de ser hospitalizado por un diagnóstico potencialmente evitable por la atención primaria es 2,5 veces mayor en Manhattan que en París (Gusmano et al., 2006), concluyendo, por tanto, que en esta última ciudad hay más acceso a una atención primaria de calidad, con capacidad resolutiva.

Una línea de avance a mi juicio prometedora es la ordenación basada en criterios de dominación estocástica bayesiana, que se ha empleado, por ejemplo, para ordenar países según la producción científica en Economía (Lubrano, 2004). El concepto es antiguo, pero en los años recientes se ha empleado para construir diferentes *rankings* de unidades productivas. Hay dominación de primer grado de una distribución de probabilidad F sobre otra distribución G si $F(x) \leq G(x) \forall x \in (0, +\infty)$, es decir, si la probabilidad de obtener al azar el valor x o menor no es menor con F que con G, independientemente del valor de x que se tome. Aplicado, por ejemplo, a la distribución de la renta de dos regiones significa que en una de ellas la renta está más concentrada que en la otra (las curvas de Lorenz respectivas no intersectan). En la ordenación de unidades productivas en vez de trabajar con distribuciones de frecuencias lo haremos con estimaciones de distribuciones de probabilidad, a partir de muestras. La dominación resuelve elegantemente el problema del solapamiento de los intervalos de confianza en estadística frecuencialista clásica.

Sumario

Las ligas de coste-efectividad

Los *ranking* de coste-efecividad comparan los costes por Año de Vida Ganados Ajustado por Calidad (AVAC) mediante intervenciones o programas de salud alternativos. Los costes se convierten, con propósitos comparativos, a unidades monetarias homogéneas (euros de 2001, por ejemplo), pero la efectividad —ganancia de salud— se puede medir de muchas maneras, dependiendo del programa que estemos evaluando. Unos salvan vidas, otros alargan la esperanza de vida, otros mejoran su calidad. Para dotarse de una métrica común, una unidad de medida de la efectividad que incorpore las dos dimensiones, cantidad y calidad de vida, se han definido los AVAC, que intentan valorar la utilidad de diferentes estados de salud y así reducir a una dimensión la salud ganada. La ratio coste-efectividad es un criterio de eficiencia que se define como lo que cuesta conseguir una mejora unitaria de salud: cuánto cuesta ganar un AVAC.

Se han publicado varias de esas listas ordenadas en diversos textos, y la Universidad de Harvard mantiene una página web con la ordenación de intervenciones sanitarias, clasificadas por problemas de salud[1]. Las ligas de coste-efectividad han gozado de la atención de los investigadores y de los políticos porque cubren una necesidad con aparente simplicidad. No obstante, sufren limitaciones que previenen contra su uso indiscriminado y simplista. No han de ser tomadas como recetas a aplicar automáticamente en las decisiones de financiar o no los tratamientos médicos y tecnologías sanitarias, sino como una orientación para no actuar a ciegas. Muchos de los problemas son similares a los de las ligas de los hospitales (necesidad de ajustes, gestión científica de la incertidumbre, contrastes de comparaciones múltiples).

Categorizamos las limitaciones de las ligas AVAC en tres tipos: a) limitaciones inherentes a cada uno de los estudios coste-utilidad y al propio instrumento AVAC (Arnesen et al., 2004; Badia et al., 1999; Bleichrodt et al., 1997; Bleichrodt et al., 2002a; Bleichrodt et al., 2002b; Dolan et al., 2005; Johannesson, 1994; Nord et al., 1999a); b) problemas relacionados con el metaanálisis; c) problemas para tomar decisiones de política sanitaria sobre la base de la información de la liga. Nos centramos en el segundo tipo de problemas, que son más estadísticos.

Las tablas que ordenan las intervenciones según coste-efectividad se elaboran mediante el metaanálisis que compara (o «agrega») resultados de múltiples estudios de evaluación realizados en diferentes espacios y tiempos. Suele haber problemas de selección de los estudios a incluir (falta de criterio científico explícito de inclusión; no se garantiza que los estudios de base hayan cumplido los estándares protocolizados de calidad de las evaluaciones económicas; se comparan estudios referidos a distintos momentos tecnológicos y a distintos países, que pueden diferir en costes y en efectividad).

También hay un serio problema de selección derivado del sesgo de publicación: es mucho más probable que se publiquen resultados favorables a los nuevos tratamientos y que los estudios financiados por la industria terminen en conclusiones positivas.

[1] http://www.hsph.harvard.edu/cearegistry

Sumario

Además, se excluyen de la liga los numerosos estudios que no emplean el AVAC como medida de utilidad, que están desigualmente distribuídos entre enfermedades, según un análisis comparativo de 455 trabajos de evaluación económica incluídos en la base de datos *Health Economic Evaluations Database* (HEED) (Anell et al., 2000).

Otra fuente de sesgo de selección es que las políticas de salud pública, basadas en intervenciones comunitarias para cambiar estilos de vida individuales, no se suelen someter a evaluación económica tanto como los tratamientos médicos individuales, animados y financiados por la industria.

Otro problema fundamental es que sólo se comparan promedios, sin considerar la incertidumbre implicada en los estudios originales ni los respectivos análisis de sensibilidad. Los intervalos de confianza de dos intervenciones pueden superponerse pero la tabla no informa al respecto. Aunque se utilice en la Medicina Basada en la Evidencia (MBE), esa evidencia es fragmentaria, incompleta e incierta. Hay incertidumbre diagnóstica, sobre la efectividad de muchos tratamientos médicos a corto y a largo plazo, y sobre los costes.

Incluso salvando las limitaciones anteriores, la ratio coste por AVAC ganado no debe ser el único criterio de decisión o priorización porque la disposición social a pagar por ganar salud depende de quiénes sean los beneficiarios y de cómo se distribuyan los AVACs ganados. Las tablas nada dicen sobre cómo se concentran las ganancias de salud y quienes son los beneficiarios. A sus efectos es lo mismo alargar 50 años la vida de una sola persona que alargar un año la vida de cada una de las 50 que sufren una enfermedad. Para la sociedad no es lo mismo (Nord et al., 1999b).

4. RANKING MULTIDIMENSIONAL. ¿CÓMO ORDENAR CUANDO LAS POLÍTICAS TIENEN MÚLTIPLES OBJETIVOS?

Generalmente las políticas tienen objetivos múltiples, complementarios o no. La cuestión es si ponderar o no ponderar. Cuanto más local sea el nivel de la gestión, se requiere mayor desagregación de indicadores y monitorizar los objetivos uno a uno. Frente a la ventaja de la simplicidad, los indicadores unidimensionales no ayudan a definir estrategias globales de regulación. Para asignar fondos entre unidades con las mismas reglas del juego, el regulador central requerirá una medida sintética de ejecución que compense las pérdidas respecto a la media de unas dimensiones con las ganancias de otras dimensiones. Cuanto más centralizado es el organismo que necesita el *ranking*, más necesidad suele haber de sintetizar las múltiples dimensiones del éxito en un único indicador global de *performance* o eficiencia organizativa. El gerente de un hospital o de una universidad, por ejemplo, necesita en su cuadro de mando indicadores separados de productividad y actividad en cada uno de los servicios y unidades, para tomar medidas específicas que resuelvan los problemas. Por ejemplo, necesitará conocer la lista de espera de cada servicio y prueba diagnóstica. En cambio la dirección regional del sistema de salud y la autoridad sanitaria central demandarán un indicador global de eficiencia.

Sumario

Los métodos que buscan el «orden global» perfecto de un grupo de unidades (sean países, hospitales o problemas de salud) son complejos y requieren ponderaciones de los diferentes *outputs* o dimensiones. Esas ponderaciones pueden ser establecidas «ex ante» por el investigador, o bien derivarse del propio análisis.

Las experiencias con ponderaciones preestablecidas

El *ranking* de sistemas de salud del mundo del informe 2000 de la OMS (WHO, 2001) resultó de la ponderación *ad hoc* (con criterios OMS) de las dimensiones que, según un panel de expertos, definen la ejecución de los sistemas de salud. Ha recibido muchas y acertadas críticas porque uniformizan con supuestos objetivos comunes a todos los países del mundo que, sin embargo, pueden tener (tienen de hecho sus propias metas distintas del resto(Williams, 2001). Confunden, pues, heterogeneidad e ineficiencia (Greene, 2003). Sin embargo, tiene el mérito de ser pionero, marcar una estela que han seguido muchos trabajos, organismos e investigadores.

En septiembre de 2001, el *Natural Health Service* (NHS) británico empezó a publicar, entre otros, un *ranking* de los consorcios que proveen asistencia hospitalaria aguda, clasificándolos con estrellas de excelencia, entre cero y tres según un amplio conjunto de indicadores de cumplimiento de objetivos clínicos, de gestión, de calidad, de tiempos de espera y otros que interesan al paciente. En 2004 el Ministerio de Salud encargó a un grupo de investigación del *Centre for Health Economics* de la Universidad de York y del *National Institute for Economic and Social Research* el diseño de una nueva metodología para medir la productividad y resultados globales del NHS. El informe, publicado en 2005 (Dawson et al., 2005), propone un índice global de *output* del NHS ajustado por calidad (el «*value weighted output index*»), que se calcularía con la siguiente fórmula:

$$I_{yt}^{xq} = \frac{\sum_j x_{jt+1} \sum_k \pi_{kt} q_{kjt+1}}{\sum_j x_{jt} \sum_k \pi_{kt} q_{kjt+1}}$$

donde x_{jt} es la cantidad del *output* j en el periodo t, q_{kjt} es la cantidad del atributo, resultado o característica k que produce la unidad j, y π_{kt} es el valor social marginal de ese atributo. El índice requiere datos de actividad (x) y de resultados (en términos de salud y satisfacción de los pacientes (q) (por ejemplo, tiempos de espera) que afectan a la utilidad o la incertidumbre asociada a esa espera. Además, es preciso tener datos de la valoración social de cada uno de esos resultados (π), que son las ponderaciones incorporadas en la fórmula.

Como para su cálculo se debiera disponer de datos inexistentes hoy por hoy en el sistema, proponen también, a modo de solución provisional, índices basados en costes (CWOI), que incorporan distintas combinaciones de cambios en supervivencia, efectos sobre la salud, tiempos de espera, satisfacción de los pacientes, reingresos y MRSA. A título ilustrativo, ofrecen resultados anuales para 1998-2004. El *output* hospitalario ajustado por calidad y valorado según costes de producción aumentó a una tasa media del 3,6% anual (3,35% sin ajustar, pag.200).

El informe británico abre nuevos horizontes y cambia perspectivas, sobre todo porque el enfoque es paciente-céntrico.

Sumario

La unidad de medida es el paciente tratado por el NHS (no el proveedor ni el comprador de los servicios); y la calidad se define en función de las dimensiones de los resultados que los pacientes valoran.

Una aproximación alternativa consiste en especificar los múltiples objetivos del sistema de salud como variables dependientes en un modelo multivariante de logros, permitiendo correlaciones entre ellos. Siguiendo este enfoque, una aplicación reciente para las autoridades sanitarias del Reino Unido (Hauck et al., 2006) emplea un modelo multinivel multiecuacional con 13 objetivos, en el que las unidades de nivel I son los distritos electorales y las de nivel II las autoridades sanitarias.

Las reglas de priorización, para ordenar distintos problemas dentro de un país, y las intervenciones políticas dirigidas a mejorarlos, constituyen otro ámbito de aplicación del «método del *ranking*». En España es notable a este respecto el proyecto de investigación sobre identificación y priorización de necesidades de salud, integrado en la red IRYSS[2].

Métodos que obtienen ponderaciones como resultado del análisis: los modelos DEA

Entre los métodos que evalúan la *eficiencia organizativa global* de unidades prestadoras de servicios homogéneos no comercializados, donde no hay precio de mercado que oriente sobre el «valor» de los bienes o servicios provistos, el Análisis Envolvente de Datos (*Data Envelopment Analysis*, DEA) tiene gran aceptación y es un punto de encuentro entre académicos y gestores. Los resultados sirven para diseñar sistemas de incentivos, pagar a proveedores o hacer un seguimiento temporal del desarrollo organizativo. Se aplican sobre todo a servicios públicos que producen múltiples *outputs* sin precio de mercado que refleje su valoración relativa: educación, sanidad, juzgados. En sanidad, se han aplicado modelos DEA para tantear la posibilidad de medir la eficiencia comparativa de autoridades sanitarias, de hospitales y de Equipos de Atención Primaria, también en España (Puig-Junoy et al., 2004). Generalmente, se plantean como métodos sustitutivos de los modelos econométricos de frontera estocástica. Al igual que estos últimos, todavía tendrán que resolver algunas cuestiones metodológicas pendientes y someterse al tribunal de los hechos para demostrar que son instrumentos útiles, precisos y robustos para orientar el reparto de fondos entre las unidades. Cómo gestionar científicamente la incertidumbre, cómo ajustar por factores «inevitables» condicionantes del entorno fuera de control de los gestores o cómo tener en cuenta la dinámica de la eficiencia son algunas de las cuestiones pendientes (Smith et al., 2005).

La ciencia todavía no garantiza que se obtengan medidas objetivas de eficiencia organizativa independientes de los instrumentos estadísticos. Hay demasiada sensibilidad de los *ranking* a pequeños cambios en el modelo o en los datos. Los resultados varían radicalmente, según se ajuste o no por factores que en principio se pueden considerar perteneciente a un entorno fuera de control para los gestores de los hospitales(Street, 2003). También suele ser difícil argumentar los resultados (justificar el *ranking* resultante con la lógi-

[2] Red de Investigación en Resultados y Servicios de Salud. http://www.rediryss.net/pub/units/rediryss/pdf/identifypriorizsns.pdf

ca de la organización, y no con la lógica matemática). Significación estadística y significación socio-política ni son sinónimos ni siempre concuerdan.

Si los efectos dinámicos son difíciles de modelizar en general, más todavía lo son en los modelos de eficiencia organizativa en sanidad. Las UTD arrastran la herencia del pasado —inercia no controlable— y toman decisiones hacia el futuro, gastando hoy para ganar mañana. La especificación dinámica de retardos y respuestas a impulsos es compleja, y requiere datos longitudinales que no siempre están disponibles o son comparables. El cambio tecnológico dificulta todavía más la comparabilidad temporal de la eficiencia organizativa.

Síntesis y conclusión

El diseño y la evaluación de las políticas públicas es objeto de una intensa atención científica. Los métodos cuantitativos (entendidos en sentido amplio) contribuyen a este fin con diversos instrumentos. La estandarización metodológica en curso responde a un nuevo paradigma, el de la Política Basada en la Evidencia (PBE), derivado por contagio de la Medicina Basada en la Evidencia (MBE). Se practica el «*Benchmarking*» tanto a nivel de mesogestión de centros como entre países, para aprender de los demás, comparar y estandarizar reformas y para asignar recursos centralizados a las unidades prestadoras de los servicios asistenciales.

En este artículo hemos presentado una visión general del uso de métodos cuantitativos para construir *ranking* que orienten las políticas públicas, con especial referencia a las de salud, y cómo emplearlos. Los métodos para obtener un *ranking* difieren según los objetivos, los datos y el modelo subyacente (grado de incertidumbre y forma de incorporar, en su caso, juicios de valor para ponderar diferentes objetivos).

Los *ranking* de criterio único incluyen tres tipos de modelos: los modelos econométricos de frontera estocástica, que miden comparativamente la eficiencia de las unidades productivas en el marco teórico de la teoría económica de la producción; los modelos de regresión logística para estimar la calidad, que se ajustan por factores de confusión y riesgos; y los *ranking* de programas públicos según coste-efectividad.

Generalmente las políticas tienen objetivos múltiples, complementarios o no. Los métodos que integran múltiples objetivos en un indicador global son más complejos que los de criterio único. La elección del método debe estar en función de su uso. Cuanto más local sea el nivel de la gestión, mayor desagregación de indicadores se requiere, obligando a monitorizar los objetivos uno a uno. Los *ranking* basados en criterios múltiples buscan el «orden global» de un grupo de unidades (sean países, departamentos públicos o problemas sociales), y requieren ponderaciones, establecidas «ex ante» por el investigador, o bien derivadas del propio análisis, como en el Análisis Envolvente de Datos (DEA). El análisis DEA es un punto de encuentro entre académicos y gestores, para diseñar sistemas de incentivos, pagar a proveedores o hacer un seguimiento temporal del desarrollo organizativo.

Estamos asistiendo a importantes avances metodológicos, incluyendo los que provienen de la estadística bayesiana y los métodos para datos de panel y para datos jerárquicos.

Sumario

En cualquier caso, debe imponerse la prudencia al leer los resultados de las «ligas», que a veces se presentan de forma excesivamente simplista y evitar hacer un uso inadecuado de la estadística, sin informar de los intervalos de confianza, o sin el necesario ajuste por riesgos. La evidencia disponible sugiere que la publicación de estos datos contribuye a mejorar la calidad de los servicios cuando tienen un nivel muy bajo. La ciencia todavía no garantiza que se obtengan medidas objetivas de eficiencia organizativa independientes de los instrumentos estadísticos. Hay demasiada sensibilidad de los *ranking* a pequeños cambios en el modelo o en los datos. Los resultados varían radicalmente, según se ajuste o no por factores que en principio se pueden considerar pertenecientes a un entorno fuera de control para los gestores. También suele ser difícil argumentar los resultados (justificar el *ranking* resultante con la lógica de la organización, y no con la lógica matemática). Significación estadística y significación socio-política ni son sinónimos ni siempre concuerdan.

REFERENCIAS BIBLIOGRÁFICAS

ANELL, A. AND NORINDER, A. (2000): Health outcome measures used in cost-effectiveness studies: a review of original articles published between 1986 and 1996. *Health Policy* 51(2):87-99.

ANGRIST, J. D. AND KRUEGER, A. B. (1998): Empirical strategies in labor economics.

ANGRIST, J. D. AND KRUEGER, A. B. (2001): *Instrumental variables and the search for identification from supply and demand to natural experiments*. Cambridge, MA: National Bureau of Economic Research.

ARNESEN, T. AND TROMMALD, M. (2004): Roughly right or precisely wrong? Systematic review of quality-of-life weights elicited with the time trade-off method. *J Health Serv.Res.Policy* 9(1):43-50.

BADIA, X., ROSET, M., AND HERDMAN, M. (1999): Inconsistent responses in three preference-elicitation methods for health states. *Soc.Sci.Med* 49(7):943-950.

BLEICHRODT, H., HERRERO, C., AND PINTO, J. L. (2002a): A proposal to solve the comparability problem in cost-utility analysis. *J Health Econ* 21(3):397-403.

BLEICHRODT, H., HERRERO, C., AND PINTO, J. L. (2002b): A proposal to solve the comparability problem in cost-utility analysis. *J Health Econ* 21(3):397-403.

BLEICHRODT, H. AND JOHANNESSON, M. (1997): Standard gamble, time trade-off and rating scale: experimental results on the *ranking* properties of QALYs. *J Health Econ* 16(2):155-175.

BROWN, A. D., GOLDACRE, M. J., HICKS, N., ROURKE, J. T., MCMURTRY, R. Y., BROWN, J. D., AND ANDERSON, G. M. (2001): Hospitalization for ambulatory care-sensitive conditions: a method for comparative access and quality studies using routinely collected statistics. *Can.J Public Health* 92(2):155-159.

CHIRIKOS, T. N. AND SEAR, A. M. (2000): Measuring hospital efficiency: a comparison of two approaches. *Health Serv.Res.* 34(6):1389-1408.

CUTLER, D. M., HUCKMAN, R. S., LANDRUM, M. B., AND NATIONAL BUREAU OF ECONOMIC RESEARCH (2004): *The role of information in medical markets an analysis of publicly reported outcomes in cardiac surgery*. Cambridge, MA: National Bureau of Economic Research.

DAWSON, D, GRAVELLE, H, O'MAHONY, M, STREET, A, WEALE, M, CASTELLI, A, JACOBS, R, KIND, P, LOVERIDGE, P, MARTIN, S, STEVENS, P, AND STOKES, L. (2005): Developing New Approaches to Measuring NHS *Outputs* and Activity. CHE Research Paper 6.

DOLAN, P., SHAW, R., TSUCHIYA, A., AND WILLIAMS, A. (2005): QALY maximisation and people's preferences: a methodological review of the literature. *Health Econ* 14(2):197-208.

FRÖLICH, M. (2004): Programme Evaluation with Multiple Treatments. *Journal of Economic Surveys* 18(2):181-224.

GARCIA-EROLES, L., ARIAS, A., AND CASAS, M. (2001): Los Top 20 2000: objetivos, ventajas y limitaciones del método. *Rev Calidad Asistencial* 16(2):107-116.

GEWEKE, J., GOWRISANKARAN, G., AND TOWN, R. J. (2001): *Bayesian inference for hospital quality in a selection model*. Cambridge, MA: National Bureau of Economic Research.

Sumario

GOLDSTEIN, H. (2003): *Multilevel statistical models*. London: E. Arnold.

GOLDSTEIN, HARVEY AND SPIEGELHALTER, DJ. (2003): League tables and their limitations: statistical issues in comparisons of institutional performance. .

GONZÁLEZ-LÓPEZ-VALCÁRCEL, B. AND BARBER, P. (1996): Changes in the efficiency of Spanish public hospitals after the introduction of Program-Contracts. *Investigaciones Económicas* XX(3):377-402.

GREEN, J. AND WINTFELD, N. (1995): Report cards on cardiac surgeons. Assessing New York State's approach. *N.Engl.J Med* 332(18):1229-1232.

GREENE, W. (2003): Distinguishing Between Heterogeneity and Inefficiency: Stochastic Frontier Analysis of the World Health Organization's Panel Data on National Health Care Systems . GUSMANO, M. K., RODWIN, V. G., AND WEISZ, D. (2006): A new way to compare health systems: avoidable hospital conditions in Manhattan and Paris. *Health Aff (Millwood.)* 25(2):510-520.

HAUCK, K AND STREET, A. Performance Assessment in the Context of Multiple Objectives: A Multivariate Multilevel Analysis (2006): Center for Health Economics University of York. HIBBARD, J. H., STOCKARD, J., AND TUSLER, M. (2003): Does publicizing hospital performance stimulate quality improvement efforts? *Health Aff (Millwood.)* 22(2):84-94.

IASIST. Top20. Benchmarks para la Excelencia 2005. http://www.iasist.com/top20/Top20_2005/Resultados/publicacion.pdf. 31-3-2006.

JACOBS, R. (2001): Alternative Methods to Examine Hospital Efficiency: Data Envelopment Analysis and Stochastic Frontier Analysis. *Health Care Management Science* 4(2):102-115.

JOHANNESSON, M. (1994): QALYs, HYEs and individual preferences—a graphical illustration. *Soc.Sci.Med* 39(12):1623-1632.

JONES, A. (2000): Health Econometrics. In A. J. Culyer and J. P. Newhouse (eds.), *North-Holland Handbook of Health Economics*. Elsevier.

KAHN III, N., AULT, T., ISENSTEIN, H., POTETZ, L., AND VAN GELDER, S. (2006): Snapshot Of Hospital Quality Reporting And Pay-For-Performance Under Medicare. *Health Affairs* 25(1):148-162.

KELLY, E. AND HURST, J. (2006): *Health Care Quality Indicators Project. Conceptual Framework Paper*. OECD Health Working Papers.

LUBRANO, M. (2004): Density inference for *ranking* European research systems in the field of economics. *Journal of Econometrics* 123(2):345-369.

NELSON, D. E., FLEMING, D. W., GRANT-WORLEY, J., AND HOUCHEN, T. (1995): Outcome-based management and public health: the Oregon Benchmarks experience. *J Public Health Manag.Pract.* 1(2):8-17.

NORD, E., PINTO, J. L., RICHARDSON, J., MENZEL, P., AND UBEL, P. (1999a): Incorporating societal concerns for fairness in numerical valuations of health programmes. *Health Econ* 8(1):25-39.

NORD, E., PINTO, J. L., RICHARDSON, J., MENZEL, P., AND UBEL, P. (1999b): Incorporating societal concerns for fairness in numerical valuations of health programmes. *Health Econ* 8(1):25-39.

OECD (2003): *A disease-Based Comparison of Health Systems. What is best and at what Cost?* Paris.

PUIG-JUNOY, J. AND ORTUN, V. (2004): Cost efficiency in primary care contracting: a stochastic frontier cost function approach. *Health Econ* 13(12):1149-1165.

SANDERSON, C. AND DIXON, J. (2000): Conditions for which onset or hospital admission is potentially preventable by timely and effective ambulatory care. *J Health Serv.Res.Policy* 5(4):222-230.

SMITH, P. AND STREET, A. (2005): Measuring the efficiency of public services: the limits of analysis. *J.R.Statist.Soc.A* 168(2):401-417.

SNIJDERS, T. A. B. AND BOSKER, R. J. (1999): *Multilevel analysisan introduction to basic and advanced multilevel modeling*. London: Sage Publications.

STREET, A. (2003). How much confidence should we place in efficiency estimates? *Health Econ* 12(11):895-907.

US DEPT.HEALTH AND HUMAN SERVICES HHS. (2006): Hospital Compare. *A quality tool for adults, including people with Medicare*. http://www.hospitalcompare.hhs.gov .

VASS, A. (2001): Doctors urge caution in interpretation of league tables. *BMJ* 323(7323):1205.

VERA-HERNANDEZ, M. (2003): [Evaluating health interventions without experiments]. *Gac.Sanit.* 17(3):238-248.

WHO (2001): *The world health report 2000 - health systems: improving performance*.

WILLIAMS, A. (2001): Science or marketing at WHO? A commentary on 'World Health 2000'. *Health Econ* 10(2):93-100.

WORTHINGTON, A. (2006): An empirical survey of frontier efficiency measurement techniques in healthcare services. http://www.bus.qut.edu.au/schools/economics/documents/disc_papers_pre2001/Worthington_67.pdf

WU, C., HANNAN, E. L., RYAN, T. J., BENNETT, E., CULLIFORD, A. T., GOLD, J. P., ISOM, O. W., JONES, R. H., MCNEIL, B., ROSE, E. A., AND SUBRAMANIAN, V. A. (2004): Is the impact of hospital and surgeon volumes on the in-hospital mortality rate for coronary artery bypass graft surgery limited to patients at high risk? *Circulation* 110(7):784-789.

Sumario

Algunas reflexiones metodológicas sobre la evaluación de la eficiencia productiva de las instituciones de educación superior

En este artículo se presentan diversas reflexiones sobre los distintos aspectos metodológicos que conlleva el proceso de evaluación de la eficiencia productiva en las instituciones de educación superior. Una peculiaridad de las instituciones universitarias es que realizan de manera simultánea dos actividades distintas, docencia e investigación, compartiendo gran parte de sus inputs. Este hecho va a estar presente a lo largo de todo el trabajo donde se repasan cuestiones de relevancia que deben ser tenidas en cuenta en todas las etapas del análisis de la eficiencia: delimitación y medición de los inputs y de los outputs de las actividades, selección del modelo y selección de la muestra.

Artikulu honetan goi-mailako ikastegien produkzio-eraginkortasuna ebaluatzeko prozesuak berekin dituen alderdi metodologikoei buruzko hainbat gogoeta jasota daude. Unibertsitateen berezitasunetako bat da bi jarduera ezberdin (irakaskuntza eta ikerketa) aldi berean egiten dituztela, eta bi jarduerok input asko partekatzen dituztela. Horixe, hain zuzen, landu da lan osoan, eta horrez gain, eraginkortasunaren azterketaren aldi guztietan kontuan hartu behar diren gai garrantzitsuak errepasatu dira: jardueren input eta output direlakoen mugatzea eta neurketa, ereduaren hautaketa eta laginaren hautaketa.

This article presents diverse reflections about the different methodological aspects that assist the process of the evaluation of the productive efficiency in the superior education institutions. A peculiarity of the university institutions is that these ones carry out simultaneously two different activities; education and research, sharing thus an important part of its inputs. This fact will be present throughout the whole work where questions of relevancy will be revised and should be taken into account during all the steps of the analysis of the efficiency: delimitation and measurement of the inputs and the outputs of the activities, type selection and sample selection.

José María Gómez Sancho
María Jesús Mancebón Torrubia
Universidad de Zaragoza

ÍNDICE

1. Introducción
2. La función de producción en educación superior
3. Técnicas de medición de la eficiencia productiva
4. La selección de la muestra
5. Conclusiones
Referencias bibliográficas

Palabras clave: evaluación, universidad, educación superior, eficiencia

N.º de clasificación JEL: D24, C44, C61, I23

1. **INTRODUCCIÓN**

La importancia de la educación superior, y de las que son sus instituciones más representativas, las universidades, queda fuera de toda duda en la sociedad actual. En España el papel y la aportación de las universidades ha sido indiscutible, sobre todo, en las últimas décadas. En un primer momento, por la extensión de sus estudios a amplias capas de la sociedad; en los últimos años, con una demanda en retroceso, se pudieron establecer otro tipo de prioridades. Entre ellas destaca el incrementar la calidad en sus actividades fundamentales (docencia e investigación).

El incremento de la calidad en las instituciones universitarias abarca, no obstante, distintas dimensiones (eficacia, excelencia, eficiencia, etc.). Una de ellas, sin duda, se relaciona directamente con la evaluación sistemática de los logros alcanzados y, en particular, de su eficiencia. Resulta difícil pensar en una universidad de calidad que sea ineficiente.

Desde un punto de vista microeconómico, el concepto de eficiencia hace referencia a los recursos empleados en la producción por una empresa, industria o, de forma general, una unidad productiva o DMU (D*ecision Making Unit*)[1]. En particular, por eficiencia productiva entenderemos la comparación entre sus valores observados y sus valores óptimos tanto para sus *inputs* como para sus *outputs*.

[1] Sobre este tema resulta clarificador el capítulo de Álvarez Pinilla (2001a).

Esta comparación puede tomar la forma de la ratio, para un nivel dado de *inputs*, del *output* obtenido respecto al máximo potencialmente obtenible; o, para un nivel dado de *outputs*, la ratio entre el mínimo potencial de *inputs* requerido y el *input* empleado; o cualquier combinación entre ambas (Lovell, 1993; pág. 4).

En los últimos veinte años se ha producido un creciente interés en el mundo académico por la medición de la eficiencia productiva en diferentes ámbitos. En particular, destacan los estudios sobre la eficiencia de diversos servicios públicos (sanidad, educación, policía, transporte, justicia, servicios públicos locales, etc.)[2].

El sector de la educación superior no ha sido ajeno a este interés por la evaluación de la eficiencia de las instituciones que lo conforman. De hecho, a fecha de hoy se dispone de un conjunto importante de trabajos dedicados al análisis de esta problemática como puede verse en el cuadro Anexo que acompaña a este artículo. En el caso español, las aplicaciones realizadas difieren en numerosos aspectos metodológicos y muestrales lo que complica la comparación de los resultados obtenidos en ellas. Entre los trabajos disponibles destacan por la minuciosidad del análisis las tesis doctorales de Giménez García (2000), Martínez Cabrera (2000), Trillo del Pozo (2002) y Gómez Sancho (2005). El primer trabajo presenta un nuevo modelo de análisis de eficiencia en costes dentro del ámbito de los modelos frontera no paramétricos, DEA y una de sus extensiones, el *Free Disposal Hull* (FDH), que se aplican para comparar los departamentos pertenecientes a la Universidad Autónoma de Barcelona[3]. Esto último le permite analizar la problemática de la eficiencia asignativa, al centrarse en los costes de los departamentos y sus posibles reducciones. En el segundo trabajo se analiza la eficiencia técnica de los departamentos de análisis económico de las universidades españolas empleando el método envolvente de datos con restricciones en las ponderaciones. En esta tesis también se evalúan dos programas concretos como son la política de creación de nuevas universidades públicas en los años ochenta y noventa, así como el programa de doctorados de calidad. En la tesis de Trillo del Pozo la atención se centra en la eficiencia técnica de los departamentos de la Universidad Politécnica de Cataluña (UPC). Para ello se aplica una batería de métodos paramétricos y no paramétricos comparando los resultados obtenidos en ellos con los que proporcionan los indicadores de gestión de la propia UPC. En el último estudio se evalúan las universidades públicas españolas utilizando una variante del método envolvente de datos, el modelo multiactividad, que permite diferenciar la eficiencia alcanzada en las funciones docente e investigadora.

En este marco se sitúa el presente artículo cuyo objetivo consiste en reflexionar sobre los distintos aspectos metodológicos que conlleva el proceso de evaluación de la eficiencia productiva de las universidades.

La estructura del trabajo se concreta en los siguientes apartados. En primer lu-

[2] Una excelente revisión de los análisis de eficiencia aplicados al sector público puede verse en Lovell y Muñiz (2003).

[3] Los modelos FDH son explicados en el manual de Cooper, Seiford y Tone (2000, págs. 105-106).

gar, se señalan las peculiaridades de la función de producción de las instituciones de educación superior y se ponen de manifiesto las dificultades asociadas a la delimitación y medición de los *inputs* y *outputs* de la actividad universitaria. En segundo lugar, se repasan brevemente las técnicas más habituales de las que se dispone para evaluar su eficiencia. En tercer lugar, se analizan los problemas relacionados con la selección de la muestra. Finalizará el presente artículo con el habitual apartado de conclusiones.

2. LA FUNCIÓN DE PRODUCCIÓN EN EDUCACIÓN SUPERIOR

La evaluación de la eficiencia productiva de cualquier organización exige un análisis del proceso productivo que se realiza en su interior.

El concepto de función de producción es clave en este contexto, ya que al indicar el máximo nivel de *output* alcanzable para cada nivel de *input*, refleja el estado actual de la tecnología en el sector analizado. Su relación con el concepto de eficiencia es directa, ya que se considera eficiente a cualquier unidad evaluada que se sitúe sobre ella e ineficiente a aquéllas que estén por debajo de dicha frontera.

A la hora de especificar la función de producción de las instituciones universitarias, resulta útil tomar como punto de partida los objetivos que habitualmente se les atribuyen[4]. Según el informe Jarrat (Committee of Vice-Chancellors and Principals of the Universities of the United Kingdom, 1985) dichos objetivos son los siguientes: instruir convenientemente a los alumnos, con el fin de dotarlos de habilidades para que puedan formar parte del mercado laboral; enseñar y promover los poderes generales de la mente; hacer avanzar el conocimiento y, finalmente, transmitir la cultura general y los principios comunes de ciudadanía. La consecución de esos objetivos se obtiene, según destaca el citado informe, a través de la investigación y la docencia, que son sus funciones principales.

La ejecución de estas funciones puede conceptualizarse de manera sencilla a través del concepto económico de función de producción, es decir, mediante la modelización de las relaciones que se establecen entre un conjunto de *inputs* o factores productivos y los *outputs* que derivan de ellos. El cuadro nº 1 permite visualizar dichas relaciones.

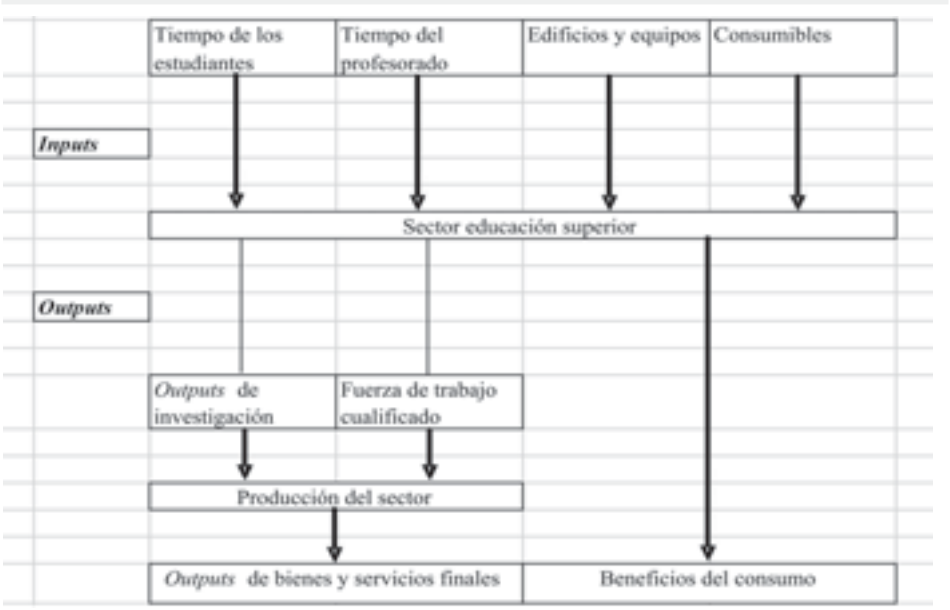
Como en él se ve, las universidades utilizan un conjunto de recursos (tiempo de los estudiantes, tiempo del profesorado, consumibles y equipos y edificios) para producir una serie de *outputs* que pueden ser clasificados, en términos generales, en productos de la docencia (conocimientos adquiridos por los estudiantes) o de la investigación (artículos, libros, patentes, etc.). A su vez, los resultados, tanto de docencia como de investigación, son un *input* intermedio en la producción de otros bienes o servicios[5]. Adicionalmente, en los resultados que genera la

[4] Un extenso estudio sobre estas cuestiones puede consultarse en las tesis doctorales de Giménez García (2000) y de Trillo del Pozo (2002).

[5] Los licenciados, output de las titulaciones, son considerados como un input para grados superiores como son los masters y los doctorados; en cuanto a la investigación, se puede pensar en los resultados de la ciencia básica como punto de partida de investigaciones aplicadas.

Cuadro Nº. 1

Proceso productivo de la educación superior



Fuente: Cave et al. (1997)

educación superior también se incluyen aquéllos que pueden proporcionar beneficios directos durante su consumo (como el dominio de una disciplina o el finalizar una investigación).

Una peculiaridad de las Universidades que se deduce del Cuadro nº 1 es que éstas realizan de manera simultánea dos actividades distintas compartiendo gran parte de sus *inputs*. En este sentido, se diferencian de otros centros educativos, dedicados exclusivamente a la actividad docente, así como de los centros de investigación, cuyos objetivos les hacen centrarse tan sólo en el proceso investigador.

Si nos detenemos a indagar en los aspectos recién mencionados se comprobará la dificultad que la conceptualización y medición de todos y cada uno de ellos genera, lo que ha imposibilitado, hasta la fecha, la especificación de una función de producción generalmente aceptada en el sector de la educación superior. En efecto, para especificar una función de este tipo se hace necesario: a) identificar y cuantificar todos los *inputs* y *outputs* relevantes y b) describir la relación entre esos *inputs* y *outputs* en términos matemáticos (Hopkins, 1990).

Respecto al primer aspecto, la identificación y cuantificación de todos los *inputs* y *outputs* relevantes, varias son las dificultades que envuelven su consecución. Quizá la más importante es la intangibilidad de muchas de sus característi-

cas, intangibilidad que afecta tanto a los *inputs* (aspectos cualitativos de alumnos, profesorado, etc.) como a los *outputs* (valor añadido del conocimiento transferido, calidad de la investigación, etc.).

En cuanto a la descripción de la relación entre *inputs* y *outputs* en términos matemáticos, la deficiente compresión que, hasta la fecha, se tiene de las tecnologías de producción subyacentes en cada una de las actividades realizadas por las instituciones de educación superior, junto al hecho antes mencionado de que nos enfrentamos a una situación de producción conjunta, complican sobremanera la expresión explícita, en términos matemáticos, de una función de producción para la educación superior[6].

Estas limitaciones relativas a la especificación de una función de producción en el ámbito de la educación superior, exige que se extremen las cautelas en la selección de las variables representativas de la actividad universitaria, en la medida en que tanto la incorporación de variables irrelevantes como la exclusión de aquéllas que contribuyen activamente en el proceso de producción pueden conducir a estimaciones de la eficiencia carentes de significación.

Una elección adecuada de las variables de referencia en una evaluación exige, por tanto, de un minucioso análisis del campo de investigación objeto de estudio. La revisión de los trabajos empíricos previos sobre la evaluación de la eficiencia de las universidades puede, sin duda, ayudar a avanzar en este conocimiento y a situar en su lugar los elementos básicos de la discusión.

2.1. Los inputs del proceso de producción de educación superior

Los recursos al servicio de las universidades suelen agruparse en las categorías de *inputs* del trabajo y del capital. Su aproximación suele efectuarse a través de indicadores sobre: a) los recursos humanos (profesorado o número de alumnos matriculados); b) los recursos financieros (gastos generales, gastos corrientes, sueldos) y c) los recursos de capital (amortizaciones, gastos en infraestructuras, superficie de las infraestructuras, etc.). Entre los problemas asociados a la cuantificación de estas variables se deben mencionar las dificultades asociadas a su estimación en unidades físicas (fundamentalmente en los recursos financieros y de capital), ya que ésta es la expresión conceptual más adecuada en un análisis de eficiencia productiva. En cualquier caso, a nuestro entender, lo más importante en relación a esta cuestión es darse cuenta de que se trata de *inputs* del proceso de educación superior que se comparten entre sus dos actividades principales. Pensemos, por ejemplo, en el profesorado o los gastos generales que se emplean simultáneamente para realizar tareas de investigación y tareas de docencia.

Variables operacionales del factor trabajo

En la mayor parte de los trabajos empíricos sobre la evaluación de las universidades el factor trabajo se ha calculado

[6] Un análisis más detallado de las complejidades que acompañan la especificación de una función de producción en el ámbito de la educación superior, tanto en su vertiente teórica como en sus hallazgos empíricos, puede verse en Hopkins (1990).

empleando la variable profesorado. Ello es lógico debido a su papel crucial en las dos actividades principales que desarrollan las universidades, tanto la docencia como la investigación. Su concreción varía en las diversas aplicaciones empíricas como puede verse en el cuadro del Anexo.

Así, una variable ampliamente empleada es el número bruto de profesores, es decir, sin ningún tipo de homogeneización. Para conseguirla, algunos trabajos diferencian por categorías, otros por dedicación, otros por áreas, o empleando diversas combinaciones de estas variables. Por último, se observa en numerosos estudios el empleo de la variable salarios del personal docente.

Los trabajos donde se especifican la diferente dedicación del profesorado (tiempo completo o tiempo parcial), suelen interpretar que el profesorado a tiempo completo realiza ambas actividades, docencia e investigación, mientras que el profesorado a tiempo parcial se dedica en exclusiva a la docencia. La realidad difiere de esta división tan categórica. No son excepcionales los casos de profesorado a tiempo completo que no investigan, ni los casos de profesorado a tiempo parcial que sí que lo hacen. Una posibilidad para integrar las figuras existentes, a nuestro juicio muy interesante ante la ausencia de datos más precisos sobre la dedicación, es recurrir al profesorado equivalente a tiempo completo que, como se comprueba en la revisión de la literatura, es la opción por la que se han decantado los trabajos revisados que tenían acceso a dicha información.

En cuanto a la discriminación en función de la categoría profesional, ya sea distinguiendo entre figuras, o por la condición de funcionario o no funcionario, se intenta incorporar un componente cualitativo a los *inputs*[7]. Una objeción a su utilización radica en que correlaciona antigüedad y calidad, hecho cuando menos cuestionable[8].

Con respecto a la utilización de los salarios consideramos que su empleo, y más en nuestro país, discrimina sólo hasta cierto punto, ya que los salarios de los funcionarios y no funcionarios de las universidades públicas están muy estandarizados, y los complementos ligados a productividad son prácticamente automáticos en el caso de los quinquenios docentes y los relacionados con los sexenios, cuya cuantía es limitada, varían mucho dependiendo del área.

Por último, la diferenciación del profesorado en función de la labor que realiza, ya sea investigadora, docente o ambas, se traduce en el caso español en considerar a los profesores asociados y a los titulares de escuela como figuras asignadas exclusivamente a la labor docente. La realidad, de nuevo, contradice lo anterior: existe una gran cantidad de profesores titulares de escuela que investigan, y

[7] La determinación de la calidad del input profesorado es mucho más compleja en educación superior que en otros niveles educativos. Ello se debe a que a los tradicionales problemas para determinar la calidad docente se añaden los de determinar su calidad investigadora. Además faltaría por contrastar qué combinación de ambos es la que define a un buen profesor universitario.

[8] Un profesor joven puede ser mejor docente y mejor investigador pero hasta que no pase determinado número de años eso no se verá reflejado en su escalafón. Por otro lado, muchas de las teorías y trabajos que han resultado más importantes para el desarrollo de los conocimientos fueron elaboradas y desarrolladas por los autores a una edad «temprana».

la figura de asociado ha sido en muchos casos empleada como mano de obra barata que empleaba a jóvenes en procesos de realización de sus tesis y, por tanto, investigaban al mismo tiempo.

Otras variables que han sido empleadas en diversos estudios sobre la evaluación de la eficiencia de las universidades de manera complementaria a la variable profesorado tienen que ver, fundamentalmente, con el personal no académico, ya sea personal de administración y servicios o la siempre difícil figura de catalogar en este tipo de estudios como son los becarios/ayudantes de investigación, los cuales tienen rasgos para poder ser considerados como alumnado o como profesorado.

Reconociendo la importancia de contar con un personal no docente adecuado en número y calidad a las necesidades de las universidades, su relación con los resultados de las actividades, docencia e investigación, es más indirecta al circunscribirse su labor, fundamentalmente, a la realización de actividades administrativas.

En definitiva, la discusión sobre la evidencia empírica previa nos lleva a considerar a la variable profesorado equivalente a tiempo completo como la más adecuada para aproximar el factor trabajo, sin que ello excluya la posibilidad de incluir otras (becarios, personal de administración, etc.), pero siempre siendo conscientes de que su inclusión disminuirá el poder discriminatorio del análisis.

Variables operacionales del factor capital

Con respecto al factor capital lo más destacable en educación superior es, como en el caso anterior, su carácter de *input* compartido por ambas actividades. En el cuadro del Anexo se sintetizan las distintas variables empleadas en los estudios empíricos para estimar este factor productivo.

Una primera diferenciación entre los trabajos radica en si han empleado indicadores físicos o monetarios, siendo los segundos los que cuentan con un mayor respaldo. Ello se debe a que la aproximación del capital a través de medidas de tipo físico no parece ser muy adecuada, ya que la relación entre este tipo de indicadores (metros cuadrados, número de ordenadores o libros en la biblioteca, etc.) y los resultados, cubierto cierto mínimo, no parece ser directa e inmediata en el sentido de que un incremento en este tipo de recursos provoque un aumento en las cantidades producidas.

Con respecto a los indicadores monetarios, éstos suelen aproximarse mediante las variables: gasto total, ingresos ligados a ayudas a la investigación o gastos de funcionamiento. El empleo de la variable gasto total presenta, a nuestro juicio, dos inconvenientes. Dentro de ella se incluye el gasto de personal con lo cual la variable profesorado estaría siendo doblemente contabilizada, como *input* de trabajo y como *input* de capital. Es decir el gasto, independientemente de la variable que se emplee, debería descontar los sueldos del profesorado si éste es incluido como *input* independiente. Por otro lado, el gasto total, al incorporar los gastos de capital, se ve muy afectado por la inversión en nueva infraestructura, lo que puede generar distorsiones en la comparación entre universidades en un año concreto.

Sumario

Como se ha señalado, otra opción por la que se han decantado algunos trabajos consiste en combinar diversos indicadores de *inputs* de capital para remarcar sus posibles usos alternativos. A nuestro entender, esto tiene sentido si lo que se pretende con ello es desligar algún tipo *input* que se emplea de manera exclusiva en una de las actividades. Uno de los más característicos son los fondos obtenidos, competitivamente, para la investigación. Esta desagregación puede resultar muy útil, como se verá más adelante, si para obtener las estimaciones se emplea el modelo multiactividad.

Por lo recién señalado, los gastos de funcionamiento, tercera opción manejada en la literatura, sobre esta materia nos parece en la actualidad la variable que mejor estima el factor capital. En todo caso, en un futuro próximo la mejora de las estadísticas universitarias pueden permitir el empleo de variables más desagregadas que recojan su empleo conjunto o separado para cada una de las actividades principales de las universidades: docencia e investigación.

2.2. **Outputs de la actividad docente y de la actividad investigadora**

La aproximación a los *outputs* de las universidades se ha llevado a cabo habitualmente a través de indicadores de carácter cuantitativo. Se distinguen habitualmente tres tipos de resultados que suelen estar relacionados con las misiones fundamentales de las instituciones universitarias. La primera de ellas es la actividad docente que representa la función de transmisión de conocimiento. La segunda es la actividad investigadora que representa la función de creación de nuevo conocimiento. En tercer lugar se encuentra la actividad de servicio público a través de servicios a la comunidad y otras actividades. Su estimación genera muchos más problemas que la de los *inputs*. El problema fundamental radica en la medición de aspectos cualitativos de las actividades docente e investigadora.

Variables operacionales de los resultados docentes

La medición del *output* docente es, a nuestro juicio, el que más dificultades genera en los análisis a nivel institucional en educación superior. No existe acuerdo unánime acerca de qué indicadores emplear, ya que entran en juego dos dimensiones: el aprendizaje en sí mismo (*output* inmediato) y la inserción laboral que deriva de éste (*output* diferido).

Partiendo de la dificultad de estimar cualquiera de ellos, nuestro interés se desplaza a hacer una valoración de los indicadores empleados en los trabajos que tratan de evaluar la eficiencia en la educación superior. Como puede verse en el cuadro situado en el Anexo, casi todos los trabajos han empleado indicadores cuantitativos. Entre los más utilizados se encuentran aquéllos que se refieren al uso del servicio (número de créditos impartidos, número de alumnos matriculados), aquéllos más relacionados con los resultados obtenidos (número de créditos aprobados, número de alumnos graduados) y aquéllos que tratan de incorporar aspectos cualitativos (empleo de encuestas docentes).

La primera opción (número de créditos impartidos y alumnos matriculados) presenta un inconveniente: con estas varia-

bles no se están midiendo resultados ni existe forma de discernir si se ha aprovechado algo esa enseñanza, ya que no se controla ni siquiera la asistencia. En cualquier caso, a nuestro juicio, el número de créditos impartidos tiene la ventaja, frente al número de matriculados, de que delimita más correctamente la carga docente que soportan las instituciones de educación superior, ya que la segunda variable no permite discriminar entre los alumnos que se matriculan en todo el curso o sólo en algunas asignaturas.

El segundo bloque de variables de *output* lo configuran diversos indicadores de resultados como son los aprobados (número de créditos superados) y los egresados o titulados. Respecto a la variable aprobados o créditos superados son variables que miden los resultados obtenidos en un año concreto, es decir, la educación aprovechada por el alumnado ese año. Entre las objeciones a su utilización se deben señalar que se trata de un resultado intermedio y que su empleo no tiene en cuenta la posibilidad de abandono antes de la consecución del título. Por el contrario, el número de titulados, tiene a su favor que sí capta los abandonos en el tramo universitario y que además aproxima el resultado final del proceso universitario: la obtención del título. Aún así debería afinarse más en su medición, ya que no tiene en cuenta el tiempo empleado por el estudiante en finalizar sus estudios. Otra debilidad reside en que es meramente cuantitativa por lo que no permite dilucidar si la enseñanza recibida en una u otra universidad es de mejor o peor calidad. En todo caso, estas variables representan mejor el *output* docente de las universidades que los indicadores de matriculados o créditos impartidos anteriormente mencionados.

Por último, el empleo de indicadores que recojan la calidad de la docencia es prácticamente inexistente; tan sólo en algunos de los trabajos realizados para evaluar a los departamentos aparece algún indicador de tipo cualitativo como son las encuestas que realizan los alumnos. Siendo éste un aspecto más del carácter multidimensional de la calidad, importante sin duda, se aleja de la evaluación del componente estrictamente académico, ya que nada asegura que de ellas se pueda extraer el valor añadido por la docencia impartida en las universidades. A ello hay que añadir la inexistencia de un diseño homogéneo entre las universidades, lo que puede dar lugar a problemas de comparación en los estudios donde se analizan distintas universidades[9].

En síntesis, la inexistencia de medidas que recojan los aspectos cualitativos de la actividad docente (tal como una prueba de evaluación de los conocimientos de carácter global por titulación, datos sobre la inserción laboral o los ingresos de los titulados, etc.) constituye realmente una de las principales problemáticas a que se enfrenta cualquier investigador sobre la eficiencia de las universidades.

---

[9] Hasta la fecha no existen datos que recojan estos aspectos cualitativos para todas las universidades españolas. Por otro lado, indicar que en el futuro sí que se podrá disponer de este tipo de indicadores cuando se ponga en marcha lo que propone el borrador del catálogo de indicadores del sistema universitario público español, en el que se recoge una batería de ellos relativos a aspectos cualitativos de la docencia como: tasa de abandono, tasa de rendimiento, tasa de éxito, tasa de graduación, duración media de los estudios, tasa de progreso normalizado, satisfacción con los estudios, satisfacción con el empleo, etc. (http: //wwwn.mec.es/ educa/ ccuniv/ html/ indicadores/ indicado.pdf).

Variables operacionales de los resultados investigadores

La investigación, junto a la docencia, constituye un binomio indisociable en el ámbito universitario. Por tanto, la medición del *output* de investigación se convierte en un paso inexcusable a la hora de realizar cualquier análisis de eficiencia productiva de las instituciones de educación superior.

La aproximación al *output* de la investigación se ha acometido, como se comprueba en el cuadro del Anexo, de tres maneras distintas: a) recurrir a los fondos de investigación, b) tratar de obtener información lo más extensa posible de los resultados de la investigación de cada universidad (artículos, libros, tesis, ponencias, patentes, etc.), o c) acudir a indicadores de calidad científica.

De todas ellas la que en el momento actual goza de un mayor respaldo es la que hace uso de indicadores de calidad científica, en particular de los factores de impacto. Es por ello por lo que en esta sección nos centraremos en valorar las ventajas e inconvenientes que presentan estos indicadores en los estudios de evaluación de la eficiencia productiva de las universidades[10].

Los índices de impacto tratan de valorar la contribución al conocimiento en función de las citas recibidas por un trabajo científico tras su publicación. A nuestro juicio, se trata de una medida objetiva al valorar la contribución de una investigación al conocimiento.

Estos índices parten del supuesto de que tendrán más impacto, y por extensión más calidad, las revistas cuyos artículos son más citados y que, en consecuencia, contribuyen en mayor medida a la generación de conocimiento científico. Son aceptados, en especial en el mundo anglosajón, como la mejor vía de medición de este concepto de calidad. El origen del empleo de esta metodología se debe a Garfield (1955) fundador del *Institute of Scientific Information*[11].

Hay que ser conscientes desde un primer momento que su empleo implica reconocer como investigación tan sólo una parte de la producción científica de las universidades (la vertida en las revistas), quedándose fuera otros formatos, importantes en determinados campos del saber, como son las patentes o los libros e incluso determinadas áreas a las que es muy difícil aplicar los factores de impacto tal y como están diseñados actualmente.

El cálculo de un factor de impacto (FI) para un año determinado se concreta en

[10] La medición de la calidad científica sigue siendo un tema complejo sobre el que no existe un consenso generalizado. Desde una perspectiva amplia pueden diferenciarse tres vías principales para llevar a cabo la medición de dicha calidad: calidad biblioteconómica, valoración de los pares (*peer review*) e índices de impacto (véase Velázquez Angona, 2002).

[11] En 1958, Eugene Garfield creó el *Institute for Scientific Information* (ISI) en Filadelfia. Este instituto comenzó su actividad modestamente confeccionando resúmenes de publicaciones de química y biomedicina. A partir de 1961, empezó a elaborar los C*itation Index*. Este Instituto ofrece las bases de datos: *Science Citation Index Expanded* que recoge información (autores, título, revista, páginas, etc.) acerca de los artículos publicados desde 1945, el *Social Science Citation Index* que recoge información desde 1956 y el *Arts & Humanities Citation Index* que recopila su información desde 1975. También desde 1975 se publica anualmente los *Journal Citation Report* de Ciencias y de Ciencias Sociales. En ellos se proporcionan los distintos indicadores de factores de impacto para la muestra de revistas seleccionadas por ISI.

dividir el número de veces que en el año considerado son citados los artículos de una revista que han sido publicados durante un periodo de tiempo determinado (dato que se extrae del vaciado de revistas del año considerado) entre los artículos publicados por ella en un periodo de tiempo dado (dato que ha sido extraído y acumulado a lo largo de los años que se consideren necesarios). Dicho de otro modo, el factor de impacto es una medida de la frecuencia con la que el «artículo promedio» de una revista ha sido citado en un año en concreto. La fórmula general sería:

$$F1\ (año\ t) = \frac{Citas\ recibidas\ en\ t\ por\ los\ artículos\ publicados\ n\ años\ antes}{Artículos\ publicados\ en\ m\ años\ antes}$$

Así, se puede obtener un FI para el propio año (n=0 y m=0) o se puede obtener un FI con un desfase determinado.

Una problemática que tiene una relevancia fundamental cuando la medición de la calidad científica se plantea en el contexto específico de la evaluación institucional hace referencia a los sesgos asociados a los factores de impacto a la hora de establecer comparaciones, bien sea entre instituciones, áreas de conocimiento o autores.

Estos tienen su origen en cuatro aspectos: a) la importancia relativa del área en la muestra de revistas seleccionadas; b) el vaciado incompleto de las revistas; c) la no consideración de las diferencias en la propensión a citar y d) la no consideración de las diferencias en el periodo de tiempo en que se materializa el impacto.

Refiriéndonos al factor de impacto más ampliamente utilizado en la actualidad, el factor de impacto del *Institute of Scientific Information* (ISI), un primer aspecto que puede dificultar la comparabilidad entre las áreas tiene que ver con la muestra de revistas seleccionadas, dado que existen áreas que no tienen ninguna representación en la base de datos (Humanidades y algunas Ciencias Sociales como Derecho). Al tratarse de un indicador multidisciplinar esto conlleva que los resultados estén sesgados a favor de las áreas que están mejor representadas o, si hablamos de subáreas, de aquéllas que estén más relacionadas con las mejor representadas.

En cuanto a los sesgos asociados al vaciado de las revistas éstos tienen su origen en que por motivos de coste ISI sólo vacía un número relativamente pequeño de ellas[12]. Esto da lugar a un problema relacionado con la autocitación: las revistas «vaciadas» al autocitarse tienen mayor impacto que las no vaciadas, por lo que para hacer una comparación correcta deberían eliminarse esas autocitas.

En cuanto a la propensión a citar, es decir, el distinto número de citas incluidas en cada artículo, tiene efectos inmediatos sobre el valor del factor de impacto analizado desde el momento en que dos áreas de conocimiento, revistas o autores tienen, por los motivos que sean, comportamientos diferenciados. En efecto, con todo lo demás idéntico entre las áreas, el hecho de que en una cite el doble que en la otra le otorgará automáticamente el doble de impacto.

[12] Ello es debido a que se ha comprobado que un número relativamente pequeño de revistas publica el núcleo sustancial de los resultados científicos significativos, en cualquier campo de conocimiento considerado. Este principio se conoce con el nombre de Ley Bradford.

Otro tipo de sesgos derivan del periodo de tiempo en que se materializa el impacto. Éstos se producen porque el factor de impacto que calcula ISI para cada año, sólo toma en consideración los datos relativos a los dos años anteriores (n=m=2), lo que penaliza a las investigaciones que, por su naturaleza, requieren un periodo de asimilación más largo. La corrección de este problema pasa, a nuestro entender, por homogeneizar el porcentaje de citas y artículos considerados en el factor de impacto de todas las revistas, en lugar de homogeneizar por años que es como lo efectúa ISI (dos años anteriores tanto para citas recibidas como para artículos publicados)[13].

En definitiva, la existencia de bases de datos internacionales como ISI permiten cualificar la investigación y así valorar la producción científica de las universidades de una manera más objetiva que las que se han empleado tradicionalmente basadas en el mero recuento de publicaciones o las que han manejado los fondos de investigación. No obstante, una correcta utilización de los factores de impacto en estudios comparativos requiere de esfuerzos por parte del investigador para minimizar todos los sesgos asociados a este tipo de indicadores y que tienen su origen en el propio proceso de construcción de los mismos.

3. TÉCNICAS DE MEDICIÓN DE LA EFICIENCIA PRODUCTIVA

Otra de las elecciones fundamentales en la evaluación de la eficiencia productiva tiene que ver con la metodología empleada para obtener las estimaciones. Las alternativas son dos: una función de producción paramétrica, estimada a través de regresiones, o bien la utilización del método no paramétrico conocido como análisis envolvente de datos (DEA), basado en técnicas de programación matemática. La clave radica en considerar las ventajas y debilidades de cada técnica en el contexto específico del sector universitario.

En este sentido, tanto los argumentos teóricos como la revisión de la literatura se inclinan por el empleo del DEA. Existen varias razones para esta decisión. La más importante, y por ello la más enunciada, hace referencia a la flexibilidad que esta técnica permite en cuanto a la modelización de la tecnología subyacente. Las posibilidades paramétricas deben asumir una forma funcional específica para la función de producción, supuesto que afecta de forma directa a los resultados de la estimación. El DEA, en cambio, permite obviar esta limitación, de tal modo que sólo se requiere que se asuman una serie de propiedades para el conjunto de posibilidades de producción. La dificultad de modelizar los procesos de producción de las dos actividades principales de las universidades —docencia e investigación— inclinan la balanza a favor de la posibilidad no paramétrica, puesto que permite aproximar de forma más fiable la tecnología de «caja negra» que subyace tanto en el proceso productivo educativo como en el investigador.

[13] Un ejemplo de correcciones para reducir los sesgos que derivan de las diferencias en la propensión a citar y en el periodo de tiempo en que se materializa el impacto puede verse en Gómez Sancho (2005).

Sumario

Además del anterior, otros argumentos apoyan el empleo del DEA en este campo. Así, la naturaleza multidimensional de los *outputs* en ambas actividades es asumida sin dificultad por este método, frente a las limitaciones de las técnicas paramétricas para tratar con múltiples *outputs*. Del mismo modo, la ausencia habitual de precios que se da en los servicios públicos (y la educación superior es un claro ejemplo) es un problema resuelto por el DEA de forma endógena, al fijar objetivamente los valores óptimos que para cada unidad evaluada maximizan su eficiencia productiva relativa. Finalmente, destacar la riqueza de los resultados ofrecidos por esta técnica de evaluación, resultados que pueden constituir una base de sumo interés en la implementación de estrategias de mejora de la gestión de los centros ineficientes[14].

Pese a las ventajas anteriores, también se deben considerar las debilidades propias del DEA con el fin de que no afecten de forma importante a los resultados de la evaluación que estamos efectuando. Entre las más destacadas cabe citar su carácter determinista, la inexistencia de un método estadístico de selección de variables y su alta sensibilidad a la presencia de *outliers* y/o errores de medida.

La evaluación de la eficiencia mediante el análisis envolvente de datos puede, no obstante, realizarse utilizando diferentes modelos de programación matemática, cada uno de los cuales, a su vez, puede aplicarse bajo versiones alternativas. De todos los modelos existentes, los que se derivan del artículo original de Charnes, Cooper y Rhodes (1978) –los denominados modelos CCR- suministran la estructura básica del método, siendo los demás extensiones de ellos que tratan de incorporar supuestos adicionales sobre la estructura de producción subyacente en los datos o sobre las características de las variables a incluir[15].

Entre todas las extensiones existentes en la literatura sobre el método envolvente, presentan un atractivo especial los trabajos de evaluación de la eficiencia de sectores —como el universitario— en los que se realizan de manera conjunta varias actividades tecnológicamente distintas utilizando recursos comunes, el modelo sugerido por Beasley (1995) y revisado posteriormente por Mar Molinero (1996) y Tsai y Mar Molinero (1998 y 2002). Es el conocido como modelo multiactividad.

La idea de partida de este modelo es que cuando una unidad productiva realiza conjuntamente varias actividades que no pueden suponerse tecnológicamente idénticas (hospitales, servicios policiales, universidades), la aplicación del modelo DEA convencional a todos sus *inputs* y *outputs* —no tiene en cuenta que la entidad no es libre de asignar todos sus *inputs* a todos sus *outputs*, ya que no todos son igualmente relevantes en las dos actividades— no reflejará la eficiencia real de la organización, a la vez que no permitirá identificar a aquellos centros que pueden ser eficientes en una de las actividades sin serlo en la otra.

[14] En Mancebón Torrubia (1998) se ofrece un análisis detallado de los resultados suministrados por el DEA y de su utilidad como instrumento de gestión pública.

[15] Una exhaustiva revisión de todos los modelos existentes puede consultarse en los manuales de Cooper, Seiford y Tone (2000), Thanassoulis (2001), Cooper, Seiford y Zhu (2004).

Para resolver este problema el modelo multiactividad plantea una modificación de la estructura matemática de los modelos convencionales en las que cada *input* y *output* específico es tratado de forma análoga a la que se da en los problemas generales, pero en la que cada una de las variables compartidas es asignada a cada actividad por el propio modelo de resolución. La formulación matemática para el caso de un sector en que se realizan dos actividades productivas (actividades A y B) con un input específico cada una (x^A y x^B) y un *input* compartido (x^C) es la siguiente:

$$Max\ \omega_0 = \theta_0^A \omega_0^A + \theta_0^B \omega_0^B$$

$$sujeto\ a:$$

$$\sum_{j=1}^{n} \lambda_j^A x_{Aj} \le x_{A0}$$

$$\sum_{j=1}^{n} \lambda_j^B x_{Bj} \le x_{B0}$$

$$\sum_{j=1}^{n} \beta \lambda_j^A x_{Cj} + \sum_{j=1}^{n} (1-\beta) \lambda_j^B x_{Cj} \le x_{C0}$$

$$\sum_{j=1}^{n} \lambda_j^A y_{Aj} \ge \omega_0^A y_{A0}$$

$$\sum_{j=1}^{n} \lambda_j^B y_{Bj} \ge \omega_0^B y_{B0}$$

$$\lambda_j^A, \lambda_j^B \ge 0$$

donde y_A el *output* de la actividad A; y_B, el *output* de la actividad B. Las variables λ^A y λ^B se corresponden con los coeficientes de la combinación lineal que da lugar a la entidad ficticia de referencia en las dos actividades analizadas (actividades A y B respectivamente). Los valores β indican la proporción del recurso compartido (x_C) asignado a la primera actividad. En cuanto a ω_0^A y ω_0^B representan, respectivamente, el inverso de las tasas de eficiencia de la entidad evaluada (DMU_0) en cada actividad (A y B) y ω_0 el inverso de la tasa de eficiencia en la actividad global realizada por la DMU_0. Los coeficientes θ_0^A y θ_0^B, por su parte, se asocian a las prioridades dadas a las diversas actividades. Finalmente, el subíndice 0 se asocia a la entidad cuya eficiencia se está evaluando.

Desde la perspectiva empírica, tras los propios trabajos en los que se planteaba esta metodología, el único estudio en el que se ha aplicado este modelo para estimar la eficiencia de las instituciones de educación superior es el de Gómez Sancho (2005). En todos los demás o se ha utilizado un modelo DEA convencional sobre el conjunto de actividades (docencia e investigación) desarrolladas por departamentos o universidades, o se ha aplicado un modelo DEA convencional a un único programa universitario, considerando las dos actividades desempeñadas por las universidades como independientes (véase el cuadro situado en el Anexo).

4. **LA SELECCIÓN DE LA MUESTRA**

El análisis envolvente de datos, desarrollado en el artículo de Charnes, Cooper y Rhodes (1978), descansa sobre la noción de eficiencia relativa introducida por Farrel en 1957. Es por ello que desde sus orígenes se ha destacado la necesidad de que las entidades analizadas sean altamente homogéneas. Dicho de otra manera: «un elemento de preferente atención es que se esté comparando lo comparable» (Albi, González-Páramo y López Casasnovas, 1997; pág. 294).

Centrándonos en el ámbito concreto de la educación superior, inmediatamente surge la duda de si la comparación re-

levante es la que atiende a la universidad en su conjunto o la que toma como unidad de observación a los departamentos o a las titulaciones, ya sea entre varias universidades o entre los (las) pertenecientes a una misma universidad.

A este respecto resulta clarificador el ejemplo que proporcionan Dyson *et al.* (2001). Según estos autores sería inapropiado comparar departamentos de ciencias con departamentos de humanidades dentro de una misma universidad, ya que los datos de gasto relativos a laboratorios, equipamiento y otros materiales similares favorecerían a los departamentos de humanidades donde dichas prácticas apenas tienen relevancia. Es decir, los *inputs* básicos del proceso de producción son mucho mayores en los primeros por lo que en una evaluación conjunta éstos tenderán a aparecer como ineficientes sin que ello sea realmente un reflejo de su gestión. El mismo argumento sería trasladable a la comparación de universidades con una oferta de titulaciones muy heterogénea.

En cualquier caso, la elección de una de las opciones de selección muestral recién señaladas debería tomar en consideración las características propias de cada unidad de observación. Así, si se opta por evaluar los departamentos, el análisis de eficiencia más relevante es el que atiende a la investigación, debido a la dificultad de imputar los resultados docentes a cada departamento. Si la opción seleccionada es la titulación, tiene más interés centrarse en los resultados docentes, ya que el profesorado, responsable de la investigación, puede estar asignado a más de una titulación o, si se evalúa en un periodo de tiempo superior al año o curso académico, haber ido cambiando de titulación. Si se opta por que la muestra sea la universidad, se supera la visión parcial de las anteriores muestras y se pueden obtener resultados de ambas actividades, docencia e investigación, teniendo en cuenta las interrelaciones y las restricciones existentes entre ellas, e incluso analizar otros aspectos como los servicios generales. La desventaja, en este caso, reside en que esa mayor generalidad puede tener como contrapartida una menor homogeneidad de la muestra.

Otra de las decisiones del investigador al plantearse evaluar la eficiencia de las universidades radica en si se deben comparar departamentos (titulaciones) entre universidades o dentro de la misma universidad. Desde la perspectiva de la gestión ambas evaluaciones resultan complementarias. La primera ofrece información sobre la posición de cada departamento (titulación) en relación al resto de departamentos (titulaciones) del área. La segunda muestra la situación relativa frente a otros departamentos (titulaciones) de la misma universidad. Si se comparan departamentos (titulaciones) entre universidades el problema de la homogeneidad se reduce sustancialmente, pero la información que se proporciona tiene menos interés para el gestor universitario, más preocupado por los departamentos (titulaciones) de su propia universidad.

En síntesis, la selección de la muestra en el ámbito de la educación superior, ofrece diversas opciones que el investigador debe valorar a la hora de llevar a cabo la misma, siendo consciente de las ventajas y debilidades de cada una de ellas.

Desde un punto de vista empírico, en la mayor parte de los trabajos sobre eva-

luación de la eficiencia de las instituciones universitarias se hace referencia al problema de la homogeneidad de la muestra bien sea para excluir observaciones o para subdividir la muestra en grupos atendiendo a algún tipo de criterio.

En concreto, los trabajos que han tratado de evaluar la eficiencia productiva de las instituciones de educación superior pueden dividirse en dos grupos atendiendo al criterio de segmentación muestral aplicado. Por un lado aquéllos que recurren a alguna clasificación externa y, por otro, aquéllos donde el propio investigador establece los criterios de segmentación que entiende que son más adecuados (véase cuadro situado en el Anexo).

Al primer grupo pertenecen los trabajos centrados en la evaluación de la eficiencia de las instituciones de educación superior de Estados Unidos y Canadá. Los estadounidenses basan la selección de la muestra en los criterios de clasificación de la *Carnegie Foundation*[16] y los canadienses siguen el criterio de división que utiliza la revista *Maclean* en su análisis anual sobre las instituciones de educación superior canadienses[17].

En la mayoría de los trabajos anteriores, aparte de realizar las segmentaciones de la muestra recién mencionadas, los autores realizan otras particiones atendiendo a distintos criterios. Así, en los trabajos sobre instituciones de educación superior estadounidenses, cuyo objetivo fundamental ha sido comparar la eficiencia de las universidades públicas con las privadas, la muestra ha sido dividida atendiendo a la naturaleza jurídica de dichas instituciones. En esos mismos trabajos se muestra como uno de los grandes problemas a la hora de homogeneizar las instituciones radica en la presencia o ausencia de facultades de medicina[18].

Otro grupo de trabajos está compuesto por investigaciones en las que es el propio autor el que segmenta la muestra para homogeneizar los datos. Los criterios de segmentación empleados son variados, en muchos casos múltiples y en algunos trabajos señalados superficialmente pero no explicados con claridad. Un primer criterio de selección de la muestra ha sido el país. Así, tenemos trabajos referidos a Gran Bretaña, Brasil, Austria, China, Australia y Sudáfrica. Otro criterio de segmentación ha consistido en suprimir parte de la muestra para garantizar la homogeneidad. Así, en algunos trabajos se excluyen aquellas instituciones que no imparten en todos los ciclos, las muy especializadas, las facultades de medicina, las menos importantes o las politécnicas. Por último, destacar el artículo de Abbott y Doucouliagos (2003) que es en el que esta cuestión es tratada con mayor profundidad, empleando para ello diversos criterios de agrupación, ya sea según la relación existente entre in-

[16] Esta clasificación existe desde 1971 y se desarrolló con el objeto de servir a la investigación en educación superior, al identificar categorías de colegios y universidades que pudieran ser homogéneos con respecto a las funciones desempeñadas, a las características de los estudiantes y al profesorado. La clasificación de las instituciones de educación superior en Estados Unidos, así como sus notas técnicas, puede localizarse en http://www.carnegiefoundation.org/Classification/.

[17] La clasificación puede consultarse en http://www.macleans.ca

[18] Las particularidades de los estudios de medicina tanto en contratación de profesorado, relación de éste con instituciones sanitarias, duración de los estudios o coste de las prácticas pone en duda su posibilidad de comparación con el resto de áreas.

vestigación y docencia, según la localización (regional o urbana), usando el análisis clúster (sin indicar sobre qué variables agrupan), o en función de su carácter de universidad tradicional o de educación avanzada.

En definitiva, en términos generales el problema de la selección de la muestra ha recibido una atención menor en los estudios sobre la evaluación de la eficiencia de las universidades. En ellos no se ha empleado ningún criterio que clasifique las universidades de una manera que podamos considerar clara y objetiva en relación a su homogeneidad. Una primera opción consistiría en acudir a algún tipo de clasificación ya elaborada. Una segunda opción consiste en emplear algún criterio empleado por otros autores. Una tercera opción consiste en proponer, siguiendo la indicación de algunos autores, una agrupación de las universidades recurriendo a una metodología objetiva de clasificación como es el análisis cluster. La utilización de este método como herramienta de homogeneización muestral en los estudios de eficiencia productiva ha sido sugerida por algunos autores (Albi, González-Páramo, y López Casasnovas, 1997; Dyson et al., 2001). Un ejemplo de esta última opción puede consultarse en el capítulo 2 de la tesis doctoral de Gómez Sancho (2005) donde se agrupan las universidades públicas empleando el análisis cluster y considerando homogéneas a aquéllas que tienen una oferta de titulaciones similar.

5. **CONCLUSIONES**

En este artículo se han planteado algunas reflexiones sobre los distintos aspectos metodológicos que conlleva el proceso de evaluación de la eficiencia productiva de las instituciones de educación superior, en concreto de las universidades.

Un primer hecho destacable es el papel crucial del investigador en este tipo de trabajos debido a la inexistencia de un marco consensuado de los procedimientos a seguir. En consonancia con ello, debe mencionarse, en primer lugar, la carencia de una teoría firme y compartida sobre los dos procesos llevados a cabo por las universidades (docencia e investigación), lo que obliga a extremar las cautelas a la hora de llevar a cabo el proceso de selección de las variables representativas de la actividad universitaria y sobre las que se ha de fundamentar la evaluación de la eficiencia. La revisión de los estudios previos sobre la función de producción de enseñanza superior y sobre la evaluación de la eficiencia de las instituciones universitarias se convierte en una referencia de gran interés a la hora de acometer esta labor.

En todos ellos se pone de manifiesto la carencia de indicadores sobre las actividades llevadas a cabo por la universidades, lo que constituye, sin duda, una limitación a la hora de llevar a cabo este tipo de evaluaciones.

En relación a la actividad docente las restricciones principales tienen que ver con la inexistencia de medidas que recojan aspectos cualitativos (conocimientos adquiridos, inserción laboral, etc.)

Con respecto al *output* investigador, su medición constituye también una cuestión muy controvertida, si bien en este caso parece que se van alcanzando unos consensos mínimos relativos a la idoneidad de utilizar factores de impacto

para cualificar la producción científica. En todo caso, como hemos tratado de poner de manifiesto, el uso de este tipo de factores debe hacerse con suma prudencia cuando el objeto de estudio es la evaluación institucional, en la medida en que su propio proceso de construcción les imprime unos sesgos que afectan seriamente a la comparación entre universidades, departamentos o áreas. Especial cautela requieren, como hemos visto, los sesgos que se derivan de las diferencias existentes entre áreas de conocimiento en relación a la propensión a citar y el periodo de materialización del impacto. Su corrección exige que se realicen ajustes previos a su utilización en los estudios de eficiencia productiva de las universidades.

Otro aspecto crucial en las fases de evaluación de la eficiencia productiva de las instituciones universitarias es, como se ha visto, la selección del modelo a emplear. Dentro de las dos grandes metodologías existentes al respecto el análisis envolvente de datos (DEA) presenta notables ventajas, ya que es el que, según gran número de autores, mejor se adapta a las múltiples particularidades del sector educativo en general y del universitario en particular. Por otra parte, dentro de las múltiples variantes matemáticas existentes dentro de esta metodología, presenta un especial interés, desde nuestro punto de vista, el modelo multiactividad propuesto inicialmente por Beasley (1995) y Mar Molinero (1996). Sus ventajas diferenciales se concretan en su capacidad para asignar objetivamente los *inputs* compartidos (profesorado, gastos, etc.) a las dos actividades realizadas por las universidades, y en que suministra tasas de eficiencia individualizadas para las actividades docente e investigadora facilitando la identificación de posibles patrones de especialización.

Por último, el trabajo ha puesto de manifiesto la relevancia de que la muestra evaluada sea altamente homogénea dado que las estimaciones obtenidas en la aplicación del análisis DEA dependen fuertemente de esa homogeneidad. En este sentido, compartimos la opinión de diversos autores que proponen la utilización de análisis cuantitativos tipo cluster para llevar a cabo la configuración de grupos de universidades comparables.

En todos los trabajos, pero fundamentalmente los desarrollados en España, queda manifiestamente claro que uno de los principales problemas, por no decir el principal, es la carencia de datos. No deja de sorprender que un sector como el universitario haya sido tan renuente a proporcionar información acerca de las actividades que realiza. Aunque en los últimos años se observa la aparición de estudios que tratan de paliar lo anterior, consideramos que nos mantenemos lejos de los mínimos que se deberían exigir a las universidades. La recopilación y difusión de indicadores constituye uno de los retos pendientes de nuestro sistema universitario.

En este sentido, resulta esperanzadora la propuesta que se contiene en el borrador del catálogo de indicadores del sistema universitario público español, encargado por el Consejo de Coordinación Universitaria, en el que se recoge la intención de recopilar una batería de ellos relativos a: oferta universitaria, demanda universitaria, recursos humanos y financieros, recursos físicos y resultados e información de contexto.

Sumario

No quisiéramos finalizar este artículo sin recalcar nuestro convencimiento de que un mejor conocimiento de las actividades que realizan nuestras universidades sólo puede redundar en su propio beneficio y en el de toda la sociedad española. No dudamos de que el acceso a nuevos datos, las mejoras continuas que se están produciendo en las técnicas y herramientas de gestión, así como la necesaria participación e implicación de todos los colectivos interesados en su mejora, especialmente de los investigadores, contribuirán, sin duda, a incrementar la calidad en todos los aspectos.

ANEXO I
ESTUDIOS EMPÍRICOS SOBRE LA EFICIENCIA EN INSTITUCIONES DE EDUCACIÓN SUPERIOR. UNIVERSIDADES

AUTOR/ES	OBJETIVO	INPUTS	OUTPUTS	HOMOGENEIDAD	MODELOS
Ahn (1987)	31 *senior colleges* y universidades en el estado de Texas en el periodo 1981-85. 161 *Doctoral institutions* en el curso 1984-85. EE.UU.	Matriculados en cursos para no graduados (ETC). Matriculados en cursos para graduados (ETC). Fondos federales para investigación y contratos.	Gastos de instrucción. Inversiones físicas. Gastos generales.	161 *Doctoral institutions* en el curso 1984-85. Separa en dos grupos las que tienen facultades de medicina y las que no; posteriormente entre públicas y privadas.	CCR y «DEA Windows».
Ahn, Charnes y Cooper (1988)	161 *Doctoral institutions* en el curso 1984-85. EE.UU.	Matriculados en cursos para no graduados (ETC). Matriculados en cursos para graduados (ETC) Fondos federales para investigación y contratos.	Gastos de instrucción. Inversiones físicas. Gastos generales.	161 *Doctoral institutions* en el curso 1984-85. Separa en dos grupos las que tienen facultades de medicina y las que no; posteriormente entre públicas y privadas.	CCR y «DEA Windows».
Rhodes y Southwick (1993)	160 *research universities* (96 públicas y 64 privadas). Tratan de determinar si son más eficientes las públicas o las privadas. EE.UU.	Número de profesores a tiempo completo. Número de profesores asociados. Número de otros profesores. Gastos anuales en mantenimiento. Gasto anual en actividades de biblioteca.	Matriculados en cursos para no graduados. Matriculados en cursos para graduados. Títulos de grado concedidos. Títulos de masters concedidos. Títulos de doctorado concedidos. Fondos para investigación obtenidos (públicos y privados, así como contratos).	Las 160 universidades seleccionadas son *research universities*.	CCR.

Sumario

AUTOR/ES	OBJETIVO	INPUTS	OUTPUTS	HOMOGENEIDAD	MODELOS
Marinho, Resende y Façanha (1997)	38 universidades federales en el año 1995. Emplea análisis de factores para agrupar inputs y outputs. Brasil.	Superficie: de edificios, hospitales y laboratorios. Número total de alumnos. Profesorado: doctor y no doctor, con y sin grado de especialización. Personal de apoyo y categoría. Gastos corrientes. Ingresos de nuevos estudiantes. Entradas de residentes médicos.	Matriculados en cursos para no graduados, masters y doctorado. Títulos concedidos (no graduados y médicos), masters y doctorados. Evaluaciones ponderadas de los cursos de master y doctorado.	38 de las 52 universidades federales brasileñas. Eliminan aquellas universidades en las que muchos de sus cursos para graduados no han sido evaluados.	BCC.
Athanassopoulos y Shale (1997)	Eficiencia en costes y eficiencia técnica en 45 universidades para el curso 1992-93. Reino Unido.	Eficiencia en costes: Gastos generales e ingresos por investigación. Eficiencia técnica: alumnos matriculados en cursos para no graduados (ETC), alumnos matriculados en cursos de postgrado (ETC), profesorado (ETC), calificación media en la prueba A-Level.	Títulos de grado concedidos. Títulos de postgrado concedidos. Evaluación institucional de la investigación.	La muestra inicial era de 52. Se excluyen aquellas que no ofrecen todos los grados y aquellas con estructuras distintas (universidades escocesas). De la Universidad de Londres sólo se escogen seis de sus grandes facultades para mantener la comparabilidad.	CCR y CCR con restricciones en las ponderaciones (eficiencia técnica).
Sarrico (1997)	24 universidades en el año 96. Reino Unido.	Calificaciones de entrada.	Ratio estudiantes profesorado, Gastos de biblioteca, Proporción de estudiantes a tiempo completo acomodados por la universidad, profesorado fijo, proporción de alumnos con mejores notas, clasificación en investigación, valor añadido, estudiantes extranjeros.	Basado *en The Times Good University Guide* se centra en instituciones que proporcionan los primeros títulos, excluyendo instituciones centradas en postgrados (London School of Economics, Royale College Art, Cranfield Institute of Technology).	Multiplicativo con rendimientos constantes y restricciones en las ponderaciones.
Hanke y Leopoldseder (1998)	11 universidades para los años 1993 y 1994. Austria.	Presupuesto excepto salarios. Salarios. Número de estudiantes.	Número de graduados. Número de horas del profesorado. Número de proyectos de investigación. Número de publicaciones.	Sólo se excluye la Universidad de Veterinaria de Viena. Se eliminan los datos de la facultad de medicina.	Multiactividad.

Sumario

AUTOR/ES	OBJETIVO	INPUTS	OUTPUTS	HOMOGENEIDAD	MODELOS
Li y Ng (2000)	84 instituciones de educación superior en el periodo 1993-95. China.	Numero de investigadores. Número de personal de apoyo. Presupuesto.	Número de manuscritos. Número de artículos. Número de reconocidos outputs de investigación. Número de contratos. Número de premios.	Análisis en un solo grupo y en tres grupos por regiones (este, central y oeste).	CCR.
Avrikan (2001)	36 universidades año 1995. Australia.	Eficiencia global: Profesorado (ETC), Personal no académico (ETC).	Número total alumnos matriculados en cursos para no graduados (ECT).Número total alumnos matriculados en cursos para graduados (ECT). Porcentaje ayudas investigación.	Un solo grupo.	CCR.
Abbot y Doucouliagos (2003)	36 universidades año 1995. Australia.	Numero total de profesorado (ETC). Número de personal no académico (ECT). Gasto total menos salarios. *Value of non-current assests.*	Número de matriculados (ETC). Matriculados en cursos para no graduados. Matriculados en cursos para graduados. Títulos de grado concedidos. Títulos de postgrado concedidos. *Research Quantum* (basado en un índice compuesto de la investigación).	Diversas agrupaciones: Un solo grupo. Separación es entre universidades urbanas y rurales. Según la ratio investigación/docencia. Comenta que se han realizado otras agrupaciones (sin especificar) empleando la metodología cluster con resultados parecidos.	CCR y BCC (ambos orientación input).
Taylor y Harris (2004)	10 universidades públicas en el periodo 1994-97. Sudáfrica.	Gasto total. Capital empleado. Capital empleado y número de estudiantes. Capital empleado y número de profesores. Capital empleado y gasto ajustado. Capital empleado y gasto total. Número de alumnos y de profesores.	Créditos obtenidos por los estudiantes graduados. Artículos en revistas aprobadas por el SAPSE (*South African Post-Secundary Education*).	Un solo grupo (excluye las 16 politécnicas y de las 21 restantes excluye 11 por insuficiencia de datos).	CCR.

AUTOR/ES	OBJETIVO	INPUTS	OUTPUTS	HOMOGENEIDAD	MODELOS
Gómez Sancho (2005)	47 universidades públicas en el año 2000. España.	Número de profesores equivalentes a tiempo completo. Gatos de funcionamiento.	Número de alumnos graduados. Valor del impacto de las publicaciones en revistas ISI en el año 2000, ponderando la adscripción a universidades de los autores y corrigiendo la propensión a citar y el periodo de tiempo en que se materializa el impacto.	Análisis cluster empleando medidas de correlación. Criterio de segmentación: oferta de titulaciones. Tres grupos de Universidades: técnicas, no técnicas y generalistas.	Multiactividad.

Fuentes: Martínez Cabrera (2000), Trillo del Pozo (2002) y elaboración propia.

ANEXO II
ESTUDIOS EMPÍRICOS SOBRE LA EFICIENCIA EN INSTITUCIONES DE EDUCACIÓN SUPERIOR. DEPARTAMENTOS

AUTOR/ES	OBJETIVO	INPUTS	OUTPUTS	MODELOS
Tomkins y Green (1988)	20 departamentos de contabilidad. Reino Unido.	Número de profesores a tiempo completo. Salarios de personal no docente. Salarios personal docente. Gastos totales.	Número de alumnos matriculados en cursos para no graduados. Número de alumnos matriculados en doctorado. Número de alumnos matriculados en masters. Ingresos totales. Publicaciones. Ingresos por investigación públicos, otros ingresos por investigación.	CCR.

Sumario

AUTOR/ES	OBJETIVO	INPUTS	OUTPUTS	MODELOS
Beasley (1990)	52 departamentos de química y 50 de física en el curso 1986-87. Reino Unido.	Gastos generales. Gastos en equipamiento. Ingresos por investigación.	Número de alumnos matriculados en cursos para no graduados. Número de alumnos matriculados en doctorado (que hacen investigación). Número de alumnos matriculados en masters (cursos postgrado docentes). Ingresos por investigación. Evaluación institucional.	CCR con restricciones en las ponderaciones.
Johnes y Johnes (1993)	36 departamentos de economía en el periodo 1994-98. Reino Unido.	Profesores dedicados a la enseñanza e investigación en personas-mes, profesores dedicados sólo a la investigación en personas-mes, fondos de investigación per cápita.	Artículos en revistas académicas, cartas en revistas académicas, libros, capítulos de libros, artículos en las principales revistas, fondos de investigación.	CCR.
Sinuay-Stern, Mehrez y Barboy (1994)	21 departamentos de la Universidad Ben-Gurion en 1988. Israel.	Salarios de los profesores, gastos operativos.	Fondos de investigación, número de publicaciones, número de alumnos graduados, número de créditos-hora impartidos por el departamento.	CCR.
Johnes y Johnes (1995)	36 departamentos de economía en el periodo 1994-98. Reino Unido.	Profesores dedicados a la enseñanza e investigación en personas-mes.	Artículos y cartas en revistas académicas. Artículos en las principales revistas académicas.	CCR con restricciones en las ponderaciones de los outputs.
Johnes (1995)	60 departamentos de economía para 1992. Reino Unido.	Profesores pagados por la universidad, profesores pagados por instituciones externas a la universidad.	Artículos en revistas académicas, libros, capítulos en libros, fondos de investigación.	BCC.
Beasley (1995)	52 departamentos de química y 50 de física en el curso 1986-87. Reino Unido.	Gastos generales. Gastos en equipamiento. Ingresos por investigación.	Número de alumnos matriculados en cursos para no graduados. Número de alumnos matriculados en doctorado. Número de alumnos matriculados en masters. Ingresos por investigación. Evaluación institucional de la investigación.	Modelo Multiactividad con restricciones en las ponderaciones.

Sumario

AUTOR/ES	OBJETIVO	INPUTS	OUTPUTS	MODELOS
Pina y Torres (1995)	22 departamentos de contabilidad en el curso 1991-92. España.	Profesores a tiempo completo, profesores a tiempo parcial, ordenadores, gasto en la adquisición de libros y revistas por profesor.	Artículos en revistas, libros, ponencias y comunicaciones a congresos, número de alumnos.	CCR.
Olesen y Petersen (1995)	18 departamentos de economía y empresa en el periodo 1975-86. Dinamarca.	Profesores. Profesores ayudantes/asistentes. Ayudantes de investigación.	Libros, artículos publicados en revistas danesas, artículos publicados en revistas extranjeras, documentos de trabajo.	CCR tiene en cuenta el ruido en los datos.
Doyle y Green (1996)	85 departamentos de empresa para el periodo 1988-92. Reino Unido.	Profesores dedicados a la investigación.	Artículos en revistas top americanas, artículos en revistas top británicas, artículos en otras revistas con referee, artículos en revistas profesionales, artículos en revistas populares, cartas, ponencias con evaluadores, ponencias sin evaluadores, libros, otras publicaciones, fondos de investigación y de no investigación, becas dirigidas y no dirigidas a investigación, alumnos de doctorado, otros alumnos de postgrado.	CCR modelo de supereficiencia. Restricción en las ponderaciones.
Sarafoglou y Haynes (1996)	14 departamentos de economía y empresa entre 1983-88. Suecia.	Profesores, otro profesorado, ayudantes de investigación.	Artículos, tesis, tesis de licenciatura, factor de impacto (SSCI).	CCR.
Madden, Savage y Kemp (1997)	29 departamentos de economía entre 1987-91. Australia.	Número de profesores.	Artículos en las principales revistas, artículos en otras revistas, libros, alumnos de licenciatura y alumnos de postgrado.	CCR.
García Valderrama y Gómez Aguilar (1999)	21 grupos de investigación de la facultad de ciencias de la Universidad de Cádiz periodo 1991-95. España.	Doctores, no doctores, catedráticos, titulares de universidad, resto personal del grupo, financiación interna, proyectos de investigación.	Número de publicaciones SCI (ponderadas o no), tesis (ponderadas o no), resto publicaciones (ponderadas o no), proyectos.	CCR y BCC.

Sumario

AUTOR/ES	OBJETIVO	INPUTS	OUTPUTS	MODELOS
Martínez Cabrera (2000, 2000 y 2003)	23 departamentos de análisis económico en el curso 1994-95. España.	Profesores numerarios, profesores no numerarios.	Artículos y capítulos en libros internacionales, artículos y capítulos en libros nacionales, libros, número de tesis doctorales.	CCR restricciones en las ponderaciones.
Giménez García (2000 y 2004)	42 departamentos de la Universitat Autònoma de Barcelona en el año 1998. España.	Gastos profesorado funcionario, gasto profesorado no funcionario, gastos funcionamiento.	Artículos publicados en revistas internacionales y nacionales entre 1995-98, libros y capítulos de libros en el mismo periodo, carga docente a través de profesores equivalentes a tiempo completo, calidad docente (encuesta a alumnos) y calidad investigación. (Tramos de investigación conseguidos dividido por tramos docentes conseguidos).	DEA y FDH.
Trillo del Pozo (2002)	34 departamentos de la Universitat Politècnica de Catalunya en el periodo 1995-99. España.	Personal docente e investigador, personal docente.	Publicaciones en revistas notables ponderadas por factor de impacto o por calidad, ingresos de investigación obtenidos a través de proyectos europeos, nacionales y regionales, indicador de calidad de la docencia (encuestas a alumnos).	CCR.

Fuentes: Martínez Cabrera (2000), Trillo del Pozo (2002) y elaboración propia.

REFERENCIAS BIBLIOGRÁFICAS

ABBOTT, M. Y DOUCOULIAGOS, C. (2003): «The efficiency of Australian universities: a data envelopment analysis», *Economics of Education Review,* 22, págs. 89-97.

AHN, T. (1987): «Efficiency and related issues in higher education: A data envelopment analysis approach», tesis doctoral, University of Texas at Austin.

AHN, T.; CHARNES, A. y COOPER, W. W. (1988): «Efficiency characterizations in different DEA models», *Socio-Economic Planning Sciencies,* 22, 6, págs. 253-257.

ALBI, E.; GONZÁLEZ-PÁRAMO, J. M. y LÓLEZ CASASNOVAS, G. (1997): *Gestión Pública. Fundamentos, técnicas y casos*., Barcelona, Ariel.

Sumario

;ALVAREZ PINILLA, A. (2001): «Concepto y medición de la eficiencia productiva» en Álvarez Pinilla, Antonio (Coordinador): *La medición de la eficiencia y la productividad,* Madrid, Pirámide, págs. 19-38.

ATHANASSOPOULOS, A. D. y SHALE, E. (1997): «Assessing the Comparative Efficiency of Higher Education Institutions in the UK by Means of Data Envelopment Analysis», *Education Economics,* 5, 2, págs. 117-134.

AVRIKAN, N. K. (2001): «Investigating technical and scale efficiencies of Australian Universities through data envelopment analysis», *Socio-Economic Planning Sciencies,* 35, págs. 57-80.

BEASLEY, J. E. (1990): «Comparing University Departments», *OMEGA - International Journal of Management Science,* 18, 2, págs. 171-183.

BEASLEY, J. E. (1995): «Determining Teaching and Research Efficiencies», *Journal of the Operational Research Society,* 46, págs. 441-452.

CAVE, M.; HANNEY, S.; HENKEL, M. y KOGAN, M. (1997): *The Use of Performance Indicators in Higher Education. The Challenge of the Quality Movement*, *Higher Education Policy Series*, Third, Great Britain, Jessica Kingsley Publishers.

CHARNES, A.; COOPER, W. W. y RHODES, E. (1978): «Measuring the efficiency of decision making units», *European Journal of Operational Research,* 2, págs. 429-444.

COMMITTEE OF VICE-CHANCELLORS AND PRINCIPALS OF THE UNIVERSITIES OF THE UNITED KINGDOM (A. Jarrat Chairman) (1985): «Report of the Steering Committee for Efficiency Studies in Universities.», Committee of Vice-Chancellors and Principals of the Universities of the United Kingdom. Steering Committee for Efficiency Studies in Universities, London, UK.

COOPER, W. W.; SEIFORD, L. M. y TONE, K. (2000): *Data Envelopment Analysis: A Comprehensive Text with Models, Applications, References and DEA-Solver Software*, United States of America, Kluwer Academic Publishers.

COOPER, W. W.; SEIFORD, L. M. y ZHU, J. (editores) (2004): *Handbook on Data Envelopment Analysis*, en Springer (Kluwer Academic Publishers) *International Series in Operations Research and Management Science*, Vol. 71, Boston, EE.UU.,

DOYLE, J. R. y GREEN, R. H. (1994): «Self and peer appraisal in higher education», *Higher Education,* 28, págs. 241-264.

DYSON, R. G.; ALLEN, R.; CAMANHO, A. S.; PODINOVSKI, V. V.; SARRICO, C. S. y SHALE, E. A. (2001): «Pitfalls and protocols in DEA», *European Journal of Operational Research,* 132, págs. 245-259.

FARRELL, M. J. (1957): «The Measurement of Productive Efficiency», *Journal of the Royal Statistical Society Series A (General),* 120, Part III, págs. 253-281.

GARCÍA VALDERRAMA, T. y GÓMEZ AGUILAR, N. (1999): «Factores determinantes de la eficiencia de los grupos de investigación en la Universidad», *Hacienda Pública Española,* 148, págs. 131-145.

GARFIELD, E. (1955): «Citation indexes for science: a new dimension in documentation through association of ideas», *Science,* 122, págs. 108-111.

GIMÉNEZ GARCÍA, V. M. (2000): «Eficiencia en costes y calidad en la universidad. Una aplicación a los departamentos de la UAB», tesis doctoral, Universitat Autónoma de Barcelona.

GÓMEZ SANCHO, J. M. (2005): «La evaluación de la eficiencia productiva de las Universidades Públicas Españolas», tesis doctoral, Universidad de Zaragoza.

HANKE, M. y LEOPOLDSERDER, T. (1998): «Comparing the Efficiency of Austrian Universities: A Data Envelopment Analysis Application», *Tertiary Education and Management,* 4, 3, págs. 191-197.

HOPKINS, D. S. P. (1990): «The Higher Education Production Function: Theoretical Foundations and Empirical Findings» en Hoenack, Stephen A. and Collins, Eileen L. (editores): *The economics of American universities: management, operations, and fiscal environment,* Albany, State University of New York Press, págs. 11-32.

JOHNES, G. (1995): «Scale and technical efficiency in the production of economic research», *Applied Economics Letters,* 2, págs. 7-11.

JOHNES, G. y JOHNES, J. (1993): «Measuring the research performance of UK economics departments: An application of data envelopment analysis», *Oxford Economic Papers,* 45, págs. 332-347.

JOHNES, J. y JOHNES, G. (1995): «Research Funding and performance in UK University Departments of Economics: A Frontier Analysis», *Economics of Education Review,* 14, 3, págs. 301-314.

LOVELL, C. A. K. (1993): «Production Frontiers and Productive Efficiency» en Fried, Harold O, Lovell, C. A. Knox, and Schmidt, Shelton S. (editores): *The Measuremente of Productive Efficiency. Techniques and Applications,* Nueva York, Oxford University Press.

LOVELL, C. A. K. y MUÑIZ PÉREZ, M. A. (2003): «Eficiencia y productividad en el sector público», *Hacienda Pública Española,* 95, págs. 47-65.

Sumario

MADDEN, G.; SAVAGE, S. y KEMP, S. (1997): «Measuring Public Sector Efficiency: A Study of economics Departments at Australian Universities», *Education Economics,* 5, 2, págs. 153-168.

MANCEBÓN TORRUBIA, M. J. (1998): «La riqueza de los resultados suministrados por un modelo envolvente de datos: una aplicación al sector de la educación secundaria», *Hacienda Pública Española,* 145, págs. 165-186.

MAR MOLINERO, C. (1996): «On the Joint Determination of Efficiencies in a Data Envelopment Analysis Context», *Journal of the Operational Research Society,* 47, págs. 1273-1279.

MARINHO, A.; RESENDE, M. y FAÇANHA, L. O. (1997): «Brazilian Federal Universities: Relative Efficiency Evaluation and Data Envelopment Analysis», *Revista Brasileira de Economia,* 51, 4, págs. 489-508.

MARTÍNEZ CABRERA, M. (2000): «La evaluación de la eficiencia técnica en las instituciones de educación superior: Una aplicación del análisis envolvente de datos», tesis doctoral, Universidad Complutense de Madrid.

MARTÍNEZ CABRERA, M. (2003): *La medición de la eficiencia en instituciones de educación superior*, Bilbao, Fundación BBVA.

NG, Y. C. y LI, S. K. (2000): «Measuring the Research Performance of Chinese Higher Education Institutions: An Application of Data Envelopment Analysis», *Education Economics,* 8, 2, págs. 139-156.

OLESEN, O. B. y PETERSEN, N. C. (1995): «Chance Constrained Efficiency Evaluation», *Management Science,* 41, 3, págs. 442-457.

PINA MARTÍNEZ, V. y TORRES PRADAS, L. (1995): «Evaluación del rendimiento de los departamentos de contabilidad de las universidades españolas», *Hacienda Pública Española,* 135, págs. 183-190.

RHODES, E. L. y SOUTHWICK, L. (1993): «Variations in public and private university efficiency», *Public Policy Applications of Management Science,* 7, págs. 145-170.

SARAFOGLOU, N. y HAYNES, K. E. (1996): «University productivity in Sweden: a demonstration and explanatory analysis for economics and business programs», *The Annals of Regional Science,* 30, págs. 285-304.

SARRICO, C. S.; HOGAN, S. M.; DYSON, R. G. y ATHANASSOPOULOS, A. D. (1997): «Data envelopment analysis and university selection», *Journal of the Operational Research Society,* 48, págs. 1163-1177.

SINUANY-STERN, Z.; MEHREZ, A. y BARBOY, A. (1994): «Academic departments efficiency via DEA», *Computers & Operations Research,* 21, 5, págs. 543-556.

TAYLOR, B. y HARRIS, G. (2004): «Relative efficiency among South African universities: A data envelopment analysis», *Higher Education,* 47, págs. 73-89.

THANASSOULIS, E. (2001): *Introduction to the Theory and Application of Data Envelopment Analysis: a foundation text with integrated software*, United States of America, Kluwer Academic Publishers.

TOMKINS, C. y GREEN, R. (1988): «An experiment in the use of data envelopment analysis for evaluating the efficiency of UK university departments of accounting», *Financial Accountability & Management,* 4, 2, págs. 147-164.

TRILLO DEL POZO, D. (2002): «La función de distancia: Un análisis de la eficiencia en la universidad», tesis doctoral, Universidad Rey Juan Carlos.

TSAI, P. F. y MAR MOLINERO, C. (1998): «The Joint Determination of Efficiencies in DEA: An application to the UK Health Service», Department of Management, University of Southampton, UK.

TSAI, P. F. y MAR MOLINERO, C. (2002): «A variable returns to scale data envelopment analysis model for the joint determination of efficiencies with an example of the UK health service», *European Journal of Operational Research,* 141, págs. 21-38.

VELÁZQUEZ ANGONA, F. J. (2002): *El impacto de las revistas científicas españolas de economía*. Programa de Estudios y Análisis 2002 de la Secretaría de Estado de Educación y Universidades del Ministerio de Educación, Cultura y Deporte.2002.

Sumario

El uso del análisis coste-beneficio en la estimación de la posibilidad de rescate de las autopistas[1]

En este trabajo se propone el uso del análisis coste-beneficio para estimar la rentabilidad social de un hipotético rescate del peaje de las autopistas en Cataluña. El análisis, que parte de los flujos de tráfico en la red viaria catalana en 2001 y de las simulaciones en años posteriores, pretende sistematizar y cuantificar todos los costes y beneficios diferenciales de transformar las autopistas de peaje directo en peaje indirecto —financiación pública—. Los resultados, sometidos a un análisis de sensibilidad, sugieren que es una medida socialmente rentable y que podría ser presupuestariamente sostenible en función del tráfico inducido.

Lan honetan kostuaren eta mozkinen analisia Kataluniako autobideetako ordain-sariaren alegiazko erreskatearen gizarte-errentagarritasuna balioesteko erabiltzea proposatu da. Analisiak 2001ean Kataluniako bide-sarean egondako zirkulazio-fluxuak eta ondorengo urteetako simulazioak hartu ditu abiapuntutzat, eta ordain-sari zuzeneko autobideak zeharkako ordain-sariko autobide bihurtzeak —finantzaketa publikoa— dituen kostu eta mozkin guztiak sistematizatu eta kuantifikatzea izan du helburu. Emaitzen gainean sentikortasun-analisia egin eta gero, aipatutako neurria gizartearentzat errentagarria dela eta, eragindako zirkulazioaren arabera, aurrekontuaren aldetik bideragarria izan litekeela ondorioztatu da.

In this article a cost-benefit analysis is used to consider the social yield of a hypothetical rescue of the freeways toll in Catalonia. The analysis, based in the traffic flows in the Catalan road network in the 2001 and in simulations for the later years, tries to systematize and quantify all the cost and profit differentials between a direct toll and indirect toll in freeways —public financing—. The results, after a sensitivity analysis, suggest it is a socially profitable measure and that could be sustainable budgetary depending on the increase on induced traffic.

[1] Este artículo es fruto del trabajo realizado para el Institut d'Estudis Territorials de Catalunya de la Universitat Pompeu Fabra: «Estudi sobre la viabilitat econòmica de la transformació del peatge directe a les autopistes en peatge indirecte» de López-Casasnovas G. y Castellanos Maduell A. Barcelona (2003).

Albert Castellanos Maduell
Guillem López-Casasnovas
Universitat Pompeu Fabra

ÍNDICE

1. Introducción y motivación del estudio.
2. El análisis coste-beneficio como instrumento de ayuda a las decisiones políticas
3. Selección de los costes y los beneficios incluidos en el acb
4. Determinación del impacto de los costes y beneficios
5. Monetización de los costes y beneficios
6. Obtención de la rentabilidad social actualizada
7. Análisis de sensibilidad de las variables
8. Conclusiones
Referencias bibliográficas

Palabras clave: autopistas, peaje, análisis coste-beneficio, rentabilidad social.

N.º de clasificación JEL: H43, R41

1. INTRODUCCIÓN Y MOTIVACIÓN DEL ESTUDIO

Existe un consenso generalizado en considerar la red viaria de alta capacidad como uno de los bienes económicos de interés colectivo de mayor trascendencia para el buen funcionamiento de la economía. En la literatura económica se encuentran numerosos estudios que identifican la existencia de redes de comunicación eficaces como una de las condiciones fundamentales para el desarrollo y el crecimiento económico.

Por consiguiente, teniendo en cuenta la presencia de externalidades positivas tan importantes en las redes de comunicación, parece suficientemente justificada la intervención del sector público. Es ampliamente conocido, sin embargo, que hay múltiples formas de instrumentar la intervención pública, y los ámbitos de financiación y gestión de la red viaria no constituyen ninguna excepción. Concretamente, se pueden clasificar las alternativas en dos grandes opciones: la primera consistiría en utilizar directamente fondos públicos para la construcción y su posterior mantenimiento; la segunda, adjudicar unas concesiones a sociedades privadas para la construcción y el mantenimiento de las vías a cambio de otorgar el derecho a cobro de un peaje directo durante un período de tiempo determinado. En la base de las diferencias entre ambas opciones se encuentra el papel ya del contribuyente, ya del usuario en su financiación, lo que remite a su vez al grado de beneficios externos derivados de dichas infraestructuras.

En el Estado español, desde 1972, se optó por seguir el modelo francés y se construyó la mayor parte de las autopistas a través del sistema de peaje directo explotadas por distintas (trece) sociedades anónimas concesionarias. La ausencia de unos criterios homogéneos en la planificación de los sistemas de gestión de la red viaria ha hecho que se hayan impuesto los criterios de rentabilidad privada sobre los de rentabilidad social: la iniciativa privada, lógicamente, sólo estará dispuesta a financiar las autopistas de peaje que sean claramente rentables ignorando aquellas que generan externalidades sociales no directamente recuperables a través de los peajes y no contempladas en su cuenta de resultados. En consecuencia, se obviaron en muchos casos criterios tan importantes como los medioambientales o los relativos a la vertebración territorial, de modo que la red de autopistas respondió básicamente a criterios de capacidad de pago y/o de demanda, y no a otros principios como el de equidad territorial.

Es conocido, sin embargo, que la existencia de un sistema de peaje directo de las autopistas crea algunas distorsiones. La más importante es probablemente el desincentivo que plantea su uso, provocando un desplazamiento del tráfico de largo recorrido de sus vías originarias (autopistas) a las que restan libres de peaje (carreteras). El resultado es un gran número de vehículos de largo recorrido circulando en condiciones de menor seguridad, que probablemente derivarán en un mayor número de accidentes de tráfico, con el alto coste social que esto comporta. Es por lo menos revelador que en los países donde se aplica el peaje directo haya un mayor número de siniestros en carretera.

Este es uno de los efectos, entre otros, de la opción seguida también en nuestro país, al margen de las implicaciones que de ella se derivan en el campo de la equidad territorial.

En un contexto en el que las administraciones públicas españolas actúan de acuerdo con criterios de rigor presupuestario, de control del déficit, resulta hoy muy difícil postular un cambio de opción financiadora desde los actuales peajes a la financiación presupuestaria general, por lo que supone plantear un aumento del gasto público sin una justificación económica robusta. Esta es precisamente la motivación del presente trabajo. Se trata para ello de valorar en un análisis coste-beneficio completo, tanto los beneficios sociales como los costes asociados a un cambio de opción, cuantificando tanto los recursos públicos destinados a compensar a las empresas concesionarias como por los efectos de la supresión del peaje directo (el aumento de recursos impositivos). El análisis se centra en la red viaria catalana, basándose en sus flujos de tráfico en el año 2001 y utilizando simulaciones de evolución para los años posteriores.

Con el fin de evaluar una alternativa, que como veremos tiene multiplicidad de efectos, no podemos basarnos en indicios, suposiciones o cuantificaciones parciales. Hace falta un marco de referencia global que avalúe todas las ventajas y los inconvenientes que comportaría la alternativa planteada en comparación con la situación actual: calcular la rentabilidad social de cada una de las opciones con el fin de vislumbrar cuál de ellas es la óp-

Sumario

tima. Para dicho propósito la evaluación económica permite el instrumento adecuado para una decisión pública informada capaz de marcar una pauta de articulación de la gestión pública.

2. EL ANÁLISIS COSTE-BENEFICIO COMO INSTRUMENTO DE AYUDA A LAS DECISIONES POLÍTICAS

El análisis coste-beneficio (ACB) constituye un marco de referencia ideal para el presente objeto de estudio. Tal y como afirma de Rus (2001), «*el análisis coste-beneficio consiste básicamente en reproducir a escala social el comportamiento racional de un individuo cuando sopesa las ventajas y desventajas de una acción no trivial*». El objetivo final es obtener el valor actual neto (VAN) de las alternativas planteadas:

$$VAN = \sum_{t=1}^{n+1} \frac{B_t}{(1+r)^{-1}} - \sum_{t=1}^{n+1} \frac{C_t}{(1+r)^{-1}}$$

donde:

B_t = beneficios del proyecto en el año t

C_t = costes del proyecto en el año t

r = tasa de descuento

n = número de años analizados

El horizonte temporal del análisis partirá de los datos disponibles del 2001, y de las previsiones realizadas hasta el 2020. Por lo tanto, el análisis coste-beneficio se realizará sobre los 18 años comprendidos entre el 2003 y el 2020. Las estimaciones de tráfico se han realizado suponiendo que el peaje se suprime en el año 2001, último año para el cual se disponían de datos reales de tráfico. La variación de los flujos de tráfico a consecuencia de una potencial supresión del peaje se han distribuido entre los primeros cinco años (2001-2005). Todas las referencias se han tomado del sistema de previsión SIMCAT a partir de los informes técnicos disponibles[2].

Siguiendo las fases clásicas de un ACB, las secciones más relevantes del estudio serán:

- Especificación del conjunto de proyectos alternativos (*2*)
- Selección de los coste y los beneficios incluidos (*3*)
- Determinación del impacto de cada uno de los costes y beneficios (*4*)
- Monetización de todos los costes y beneficios (*5*)
- Obtención de la rentabilidad social actualizada (*6*)
- Análisis de sensibilidad (7)

Las alternativas analizadas son:

A. Eliminación del peaje directo y sustitución por un sistema de peaje indirecto.

B. Mantenimiento del *status quo* (sistema de peaje directo actual)

[2]Todas las estimaciones y cálculos referentes a la evolución del tráfico y sus implicaciones (accidentabilidad, consumo de combustible, tiempo de trayecto,...) están basadas en las estimaciones realizadas por la empresa de sistemas de soporte a la decisión MCRIT, a partir de la aplicación del sistema de previsión SIMCAT (http://www.mcrit.com).

Cabría también incorporar alternativas intermedias, que se centrarían básicamente en la gradualidad con la cual se sustituye el peaje directo por el peaje indirecto.

Con el fin de ilustrar los resultados del análisis, se presentarán a continuación los beneficios y los costes de la supresión del peaje directo (opción A) de forma diferencial[3] respecto de lo que supondría el mantenimiento de la situación actual (opción B). En consecuencia, la conveniencia y/o optimalidad de la alternativa A implicará que VAN (A) > 0.

3. SELECCIÓN DE LOS COSTES Y LOS BENEFICIOS INCLUIDOS EN EL ACB

En primer lugar, se valora el conjunto de beneficios que aumentan el excedente de los consumidores (usuarios de la red viaria). Así:

1) Reducción del coste del trayecto: los usuarios actuales de las autopistas utilizarán el mismo servicio a un menor coste:

1.a Coste monetario: lo que dejan de pagar en concepto de peajes

1.b Coste del tiempo: La supresión del peaje directo también implica una reducción del tiempo de trayecto para los nuevos usuarios de las autopistas. Dentro de este apartado sería conveniente no obviar los dos factores siguientes:

— contabilización de la reducción de congestiones en las posibles vías alternativas a las autopistas de peaje,

— los costes de posibles congestiones que se pudieran producir en la misma autopista como consecuencia de un aumento en su utilización en ausencia de peaje directo.

En lo que respecta a las administraciones públicas:

2) Aumento de recaudación impositiva derivada de:

2.a. El impuesto sobre los carburantes: al postular que la supresión del peaje directo supone un aumento global del tráfico y, en consecuencia, un aumento de la recaudación del impuesto sobre carburantes.

2.b. Mayor recaudación impositiva adicional: al asumir que el mayor aprovechamiento de la red viaria causado por la supresión del peaje directo derivaría en un probable aumento de la actividad económica; y que, a su vez, esta mayor actividad económica genera una mayor recaudación será necesario estimar para ello la magnitud de este posible impacto recaudatorio, que en principio se prevé positivo.

En el campo de las externalidades sociales, se pueden derivar algunos de los beneficios de mayor importancia:

3) Aumento de la seguridad vial. El aumento de seguridad vial comporta la asunción de diversas hipótesis previas:

— La supresión del peaje generará dos movimientos de tráfico: vehículos que pasarán de vías alternativas a circular por las autopistas; y nuevos vehículos que no circulaban anteriormente por la red viaria (tráfico inducido) pasarán a hacerlo.

— La siniestralidad /accidentabilidad de las autopistas se mantiene a los

[3] Variaciones en positivo y negativo del excedente social.

niveles inferiores a la de las vías alternativas.

— La reducción de siniestralidad resultante de los conductores que pasan a utilizar vías más seguras (autopistas) es superior al aumento generado por la mayor circulación generado por el tráfico inducido anteriormente comentado.

Los costes potenciales:

4) Recursos que deben aportar las administraciones públicas (AA.PP.) en compensación a las empresas concesionarias de las autopistas de peaje con el fin de cubrir:

4.a Los costes de explotación: ya que probablemente, con la supresión del peaje éstos aumentarán por la mayor utilización de las vías; aunque a la par disminuyan los costes de administración por la supresión del sistema de recaudación de los peajes directos

4.b La dotación del fondo de reversión

4.c Los costes financieros netos; y finalmente

4.d El beneficio de las empresas concesionarias. En este punto cabe interrogarse sobre la posibilidad que las concesionarias puedan renunciar a un porcentaje de sus beneficios, obteniendo una menor rentabilidad, a cambio de una mayor seguridad (por ejemplo ante una creciente presión popular a favor de su abolición o de la posibilidad de la aparición de vías alternativas a las actuales realmente competitivas, impulsadas por las AA.PP. a costes inferiores, autovías, etc...).

5) Impuestos no recaudados: Impuestos que se dejarían de recaudar como consecuencia de la desaparición del peaje directo; entre éstos:

5.a El IVA aplicado sobre el mismo peaje

5.b El impuesto sobre sociedades que dejaría de recaudar el Estado puesto que las empresas concesionarias dejarían de obtener un beneficio directo. Cualquier beneficio (hecho imponible) que puedan obtener las empresas concesionarias sería el resultado de una transferencia provinente de las mismas AA.PP. que recaudan el impuesto[4].

Por último, como externalidad social negativa, debemos considerar:

6) El impacto ambiental: el aumento de la circulación derivado de la supresión del peaje puede generar un aumento de la emisión de gases contaminantes con un impacto negativo sobre el medioambiente.

En el cuadro n.º 1 aparece resumido el conjunto de costes y beneficios que tomaremos en consideración en nuestro análisis.

Cabe finalmente apuntar que aunque sean efectos muy difícilmente cuantificables (y por lo tanto no los incluyamos en el cálculo), no debemos menospreciar que la supresión del peaje aporta dos beneficios asociados:

— Aumento de la equidad territorial. El mantenimiento del peaje supone una vulneración evidente de los principios de equidad que discrimina de

[4] A excepción de las actividades complementarias de las empresas concesionarias: gasolineras, telecomunicaciones, párkings, etc.

Cuadro n.º 1

Costes y beneficios de la supresión del peaje directo

	Beneficios	Costes
Consumidores (usuarios de la red viaria)		
	↓ 1.a Coste monetario	
	↓ 1.b Tiempo trayecto	
Administraciones públicas		
	↑ 2.a Imp. carburantes	↓ 5.a IVA
	↑ 2.b Imp. adicionales	↓ 5.b Imp. Sociedades
		4.Compensación empresas concesionarias
Externalidades sociales		
	↑ 3. Seguridad vial[1]	↑ 6. Contaminación
	↑ Equidad territorial	
	↑ Uso racional territorio	

Fuente: Elaboración propia
Notas: (1) No todo el aumento de la seguridad vial representa una externalidad positiva. A nivel individual, la reducción del riesgo de accidentabilidad supone también un aumento del excedente del consumidor.

forma ostensible algunos territorios respecto otros[5].

— Un uso más racional del territorio. El peaje directo, a causa de su heterogeneidad, provoca distorsiones notables respecto de otros costes de desplazamiento en una situación de libre circulación (en ausencia de peajes directos), con las implicaciones en el uso del territorio que esto puede comportar.

4. DETERMINACIÓN DEL IMPACTO DE LOS COSTES Y BENEFICIOS

En este apartado, además de especificar el método de cuantificación de cada una de las variables, se determinará su valor en el escenario inicial para, posteriormente, analizar escenarios alternativos. Al cuantificar cada uno de los costes y beneficios, nos encontramos con cuatro unidades de medida distintas:

— Monetaria: Impuestos (2 y 5), reducción del coste monetario (1.a) y compensaciones (4)

— Siniestros/accidentes: aumento de la seguridad vial (3)

[5] Cataluña concentra el 29% de los km de autopistas de peaje directo de todo el Estado.

— Temporal: ahorro de tiempo de trayecto (1.b)

— Indicadores de polución: aumento de la contaminación atmosférica (6)

4.1. **Método de cuantificación**

1.a Reducción del coste monetario (RCM)

En primer lugar, debemos calcular el monto total que dejan de pagar los usuarios de las autopistas en forma de peaje, ya que éste será uno de los componentes del aumento de excedente del consumidor. La cantidad que dejarán de pagar los usuarios en el año t será igual a la recaudación esperada (RE_t) en peajes en el mismo año suponiendo que se mantuviera la situación actual:

$$RCM_t = RE(B)_t$$

Para calcular la recaudación esperada, proyectaremos la recaudación del año 2002[6] (R_{2002}) al año t, en función de las previsiones de evolución del volumen de tráfico (veh-km) en las autopistas de peaje suponiendo que se mantuviera el sistema de peaje directo[7].

$$RE_t = R_{2002} \cdot \frac{TAP(B)_t}{TAP_{2002}}$$

siendo:

$TAP(B)_t$ = el volumen de tráfico diario (veh-km) en las autopistas de peaje en el año t en el supuesto que se mantuviera el sistema de peaje directo.

1.b Reducción de tiempo de trayecto

Para simplificar el cálculo y trabajar con los datos disponibles, se ha utilizado una estimación de ahorro del tiempo total de desplazamiento para el conjunto de la red viaria que supondría la supresión del peaje directo a partir de los datos de tiempo medio del trayecto en las horas punta (de máximo volumen de tráfico)[8].

El resultado será una aproximación a la totalidad de tiempo que se ahorra el conjunto de usuarios que ya utilizaba la red viaria con el sistema de peaje directo. Se ha proyectado a todas las horas del día el tiempo ahorrado que se generaría en la hora punta. Un análisis de sensibilidad resultará aquí apropiado.

2.a. Impuesto sobre los carburantes

La recaudación del impuesto sobre carburantes dependerá del volumen total de tráfico en la red viaria y de los tipos impositivos aplicados al consumo. Para ello se ha distiguido el consumo de carburante entre tres tipos de combustible a los que se aplica un tipo diferenciado[8].

El aumento de consumo de carburante se estima[8] en un 5% para el año 2001 en el caso que se suprimiera el peaje directo. Los tipos impositivos necesarios para realizar los cálculos se obtuvieron de la

[6] A pesar de que el trabajo se base en datos de flujos de tráfico del 2001, se consideró más adecuado utilizar los datos de recaudación impositiva disponibles más actuales en el momento de realizar el trabajo (del año 2002) puesto que inciden multitud de factores no relacionados directamente con el volumen de tráfico.

[7] Ver punto 1.a del apartado 3.

[8] Estimaciones facilitadas por el centro de Sistemas de Soporte a la decisión MCRIT (http://www.mcrit.com/).

modificación de la Ley 38/1992[9] correspondientes al 2003. Para calcular el porcentaje de incremento que corresponde a las diferentes tipologías de consumo, se recurrió a los últimos datos disponibles de consumo de los diferentes carburantes de automoción. Con estos datos, se realizó una estimación del aumento de recaudación del impuesto de hidrocarburos para cada uno de los t años analizados.

Debemos tener en cuenta, también, que el aumento del consumo de carburante previsiblemente generará un aumento de la recaudación de IVA. En este caso se supone un tipo general del 16% sobre las diferentes clases de carburante, pero se aplica sobre un precio diferente. En consecuencia, se estiman también los diferentes porcentajes de consumo y los precios de los diferentes carburantes en Cataluña[10], considerando que casi la totalidad de consumo es de gasolina convencional.

2.b. Recaudación impositiva adicional[11]

En este apartado se pretende hacer una previsión de la recaudación impositiva global como consecuencia del aumento de la movilidad. Estimar este aumento global de recaudación requiere un proceso en dos etapas:

— Establecer una relación entre el nivel de circulación en la red viaria catalana y la actividad económica. Determinar esta relación causa-efecto es probablemente el paso más complejo ($\varepsilon_{TXV,PIB}$).

— Encontrar un indicador de elasticidad-renta de la recaudación impositiva ($\varepsilon_{PIB,T}$).

Por lo tanto, los efectos de la supresión del peaje directo en concepto de recaudación impositiva general en el año t (T_t) se pueden estimar a partir del porcentaje de aumento de tránsito estimado en la red viaria catalana si se eliminara el sistema de pago directo:

$$\Delta\%T_t = \left(\left(\frac{TXV(A)_t - TXV(B)_t}{TXV(B)_t}\right) \cdot \varepsilon_{TXV,PIB}\right) \cdot \varepsilon_{PIB,T}$$

siendo:

$\Delta\%T$ = Aumento porcentual de la recaudación impositiva.

$TXV(A)_t$ = Volumen de tráfico diario (veh-km) en la red viaria catalana en el año t en el supuesto que se suprimiera el peaje directo[12].

$TXV(B)_t$ = Volumen de tráfico diario (veh-km) en la red viaria catalana en el año t en el supuesto que se mantuviera el peaje directo[12].

No es posible encontrar una cifra específica para Cataluña de la elasticidad entre el volumen de tráfico y la actividad económica ($\varepsilon_{TXV,PIB}$) en la literatura[13]. La referen-

[9] Artículo 50 de la Ley 38/1992, de 28 de diciembre de Impuestos Especiales: http://documentacion.minhac.es/doc/Tributos/Impuestos%20Especiales/Ley_38-1992.pdf

[10] Ministerio de Economía. Informe mensual sobre precios de carburantes (agosto 2003): http://www.mineco.es/energia/hidrocarburos/pmpc/2003/junio2003.pdf

[11] Se pretende calcular el aumento de recaudación adicional, calculando el aumento de recaudación global y restarle el aumento de recaudación estrictamente provinente de los hidrocarburos (2.a).

[12] Simulaciones de tráfico diario realizadas por MCRIT.

[13] Para estimar el efecto real del tráfico sobre el PIB, se requeriría un estudio econométrico de una

cia conceptualmente más próxima se encuentra en un estudio de Matas y Raymond (1999), que precisamente analiza la elasticidad de demanda de las autopistas de peaje. Una de las conclusiones del estudio, que comprende todas las autopistas de peaje españolas desde 1980 a 1998, es que el tráfico en las autopistas de peaje está fuertemente correlacionado con el nivel de actividad económica. El estudio diferencia entre la elasticidad a corto plazo (con un valor de 0,889) y largo plazo (1,440). La obtención de un indicador de sensibilidad del PIB respecto al tráfico en las autopistas de peaje no podemos limitarlo simplemente a través de invertir estas cifras (elasticidad inversa), ya que supondría obviar otros múltiples efectos que explicarían un aumento de la actividad económica.

Para encontrar un indicador de elasticidad de la recaudación impositiva ($\varepsilon_{PIB,T}$), podemos recurrir a un estudio de Raymond y González-Páramo (1988), que ofrece uno de los trabajos más completos de elasticidad de la recaudación impositiva en España a partir de la serie temporal analizada (1955-1986). La elasticidad de recaudación de impuestos respecto el PIB se estima en 1,125. Si realizamos un cálculo más simple[14] con datos más actuales (1986-2001) para Cataluña, el valor de la elasticidad estimada es próximo al anterior: 1,208.

Estos resultados no se han computado como beneficios al cálculo global de la evaluación ya que su cuantificación requiere un conjunto de supuestos que, aunque plausibles, no son concretables en un cálculo simple.

3. Beneficios del aumento de la seguridad vial

Los efectos sobre la seguridad vial por la supresión del peaje son tres:

— Aumento de la seguridad vial por eliminación de la confusión de tráficos de largo y de corto recorrido, tanto en las autopistas como en las carreteras, al suprimir el peaje directo.

— Aumento de la seguridad vial como consecuencia de los nuevos usuarios de las autopistas que antes utilizaban rutas alternativas para evitar el pago de peaje (antiguos usuarios de vías alternativas).

— Aumento de los accidentes como consecuencia del aumento general del tráfico de usuarios en las autopistas que antes no utilizaban ninguna otra vía alternativa (nuevos usuarios).

Como fase previa a la monetización de los valores, debemos diferenciar:

— El número de accidentes de tráfico

— El número de víctimas mortales

— El número de heridos graves

— El número de heridos leves

4. Compensación a las empresas concesionarias

Para estimar las compensaciones que, para no afectar a la seguridad jurídica, se tendrían que ofrecer a las N empresas concesionarias, debemos tener en cuenta que la supresión de los peajes directos modifica la estructura de costes de estas empresas[15]. Por un lado, hay, en general,

magnitud tan amplia como el número de factores que pueden explicar el nivel de actividad económica.

[14] PIB en valores constantes como única variable explicativa de la recaudación en valores constantes.

[15] Aunque estos efectos puedan diferir entre empresas, se analizarán conjuntamente las empresas concesionarias que operan en Cataluña.

costes que varían en función al volumen de circulación en las autopistas de peaje, por otro lado, se deben considerar los elementos de negociación y reducción del riesgo presentes en la compensación de los beneficios empresariales.

4.a Costes de explotación

Entre los costes de explotación debemos diferenciar entre los costes de personal y el resto de costes de explotación. Por lo que respecta a los costes de personal, distinguimos entre tres tipos distintos, con sus correspondientes índices de evolución:

- Personal de cobro de peaje directo: asumimos que este coste desaparece a partir del 2003 bajo el supuesto de desaparición del peaje directo.
- Personal de mantenimiento: asumimos que el aumento del coste es proporcional al aumento de circulación en las autopistas de peaje bajo el supuesto de supresión.
- Personal de servicios generales: asumimos que el coste real es constante.

Hemos supuesto que el resto de costes de explotación evolucionan proporcionalmente a los costes de personal una vez ya se han suprimido los peajes.

4.b Dotación del fondo de reversión y dotaciones para la amortización de inmovilizado

Las sociedades concesionarias tienen la obligación de crear, durante el período de concesión, un fondo de reversión que iguale el total de activos revertibles a la administración que los ha concedido. La dotación del fondo de reversión será la misma, independientemente de la supresión o no del peaje. Por lo tanto, se considera que representa un valor real constante a lo largo de los t años.

4.c Costes financieros netos

De manera similar al fondo de reversión, se considera que los costes financieros son independientes de la decisión de suprimir el sistema de peaje directo. Por consiguiente, supondremos que estos costes varían de acuerdo con la evolución esperada de los ingresos.

4.d Beneficio neto de las empresas concesionarias

Hay ciertos elementos que nos indican que sería más conveniente tratar esta partida de forma diferenciada. Desde el punto de vista de las empresas concesionarias, el mantenimiento de los beneficios a largo plazo, en la situación actual, genera cierta incertidumbre que amenaza la estabilidad del negocio. Una decisión política de construir por parte de las administraciones públicas (como permiten las normas de las concesiones hoy vigentes) vías alternativas libres de peaje, afectaría seriamente al negocio.

Si asumimos que los agentes son adversos al riesgo, aún siendo imposible cuantificar dicha aversión, se proponen tres escenarios alternativos a partir de la estimación de los beneficios netos esperados. Así, en la evolución de los beneficios netos durante los últimos años, resulta posible incluir: (1) el 100% de los beneficios netos esperados en futuras compensaciones; (2) el 75% de los beneficios netos esperados en futuras compensaciones, o (3) el 50% de los beneficios netos esperados en futuras compensaciones.

El valor total actualizado a 2003 de las compensaciones a todas las concesionarias hasta el 2020 (comprendido entre

unos márgenes mínimos y máximos) es de magnitud similar a otras estimaciones realizadas bajo la hipótesis de rescatar todas las concesiones existentes en Cataluña (Serra 1999)[16], lo que otorga mayor verosimilitud a los cálculos.

5.a. Desaparición del IVA aplicado sobre el peaje directo

La supresión del peaje directo implica que el IVA aplicado también desaparece, repercutiendo en una disminución de la recaudación total. Los tipos impositivos son de un 7% para la mayor parte de vehículos ligeros, y de un 16% para los pesados. Por lo tanto, tendremos que aplicar un tipo diferenciado en función de la proporción de tráfico (veh-km) esperado de vehículos pesados y ligeros en cada año t, que se aplicará a la evolución de la recaudación en los tramos de las autopistas de peaje. Se ha supuesto que el valor real de los peajes no variaría en el caso de que se mantuviera el sistema actual.

5.b. Desaparición del impuesto de sociedades sobre las empresas concesionarias[17]

La base para estimar los impuestos que las AA.PP. dejan de recaudar durante el período analizado es la cuantía que han pagado estas empresas durante los últimos ejercicios. En el caso del impuesto de sociedades, el método más directo es proyectar los beneficios del peaje (hecho imponible) hasta el 2020 en función de la evolución reciente[18].

6. Impacto ambiental

Como variable del impacto ambiental resultante del aumento de la circulación en la red viaria catalana, siguiendo el procedimiento utilizado en el *Green Book*[19], se considera únicamente el derivado de las emisiones de CO2, cuya evolución se estima directamente proporcional al consumo de combustible. No se consideran los continuos avances en la disminución unitaria de emisión de gases contaminantes con motores más eficientes. Partiendo de las previsiones de aumento de consumo de carburante, la formulación es:

$$CO2_t = \lambda \cdot AC_t$$

siendo:

$CO2_t$ las emisiones de CO2 en el año t (toneladas),

AC_t el aumento de consumo de carburante en el año t,

λ el factor de transformación equivalente a 2,6 kg de CO2 por cada litro de carburante consumido[20].

5. MONETIZACIÓN DE LOS COSTES Y BENEFICIOS

Las variables no inicialmente expresadas en unidades monetarias son:

[16] La horquilla del valor total del rescate calculada por Serra (1999) se sitúa entre 500 mil millones y 1 billón de pesetas corrientes, mientras que los valores totales de las compensaciones hasta el 2020 estimadas en este estudio no llegan a los 4.300 millones de euros, aproximadamente 650 mil millones de pesetas de 1999.

[17] Así como de otros tributos ligados a la recaudación por peaje aplicados sobre las empresas concesionarias.

[18] Informes anuales de las empresas concesionarias

[19] Treasury HM (2003)

[20] Datos facilitados por Mcrit (http://www.mcrit.com)

— Reducción del tiempo de trayecto (1.b)

— Aumento de la seguridad vial (3)

— Impacto ambiental (6)

En general, un proceso de monetización implica asignar valores monetarios a productos que no están en el mercado, sea porque son bienes no comercializables o porque son simplemente bienes intangibles. El proceso implica asignar un valor a través de la identificación de su coste o bien de la disponibilidad a pagar (*willingness to pay*) por parte de los agentes.

1.b. Reducción del tiempo de trayecto

La opción más directa consiste en calcular la reducción del coste de oportunidad resultante del ahorro de tiempo. Para cuantificar este coste de oportunidad optaremos por el método que aplica Riera (1993), quien considera que el ahorro de tiempo tiene un valor diferente en función del motivo del desplazamiento.

Por lo tanto, un primer paso es, a través de la encuesta de movilidad cotidiana[21] (EMC) del 2001, determinar el porcentaje de viajes de movilidad no obligada (28%), y los de movilidad obligada (24,3%) (motivos profesionales o de estudio) o bien de vuelta a casa[22] (47,7%). Asignamos un valor del tiempo igual al salario medio por hora a los viajes por movilidad obligada. El valor del tiempo para la movilidad no obligada la obtuvimos de los últimos valores publicados por la *Autoritat del Transport Metropolità*[23]: el valor monetario por hora estimado a valores del 2003, es de 12,53 euros. Los valores asignados al ahorro de tiempo en años posteriores se han proyectado en función de la evolución de la productividad anual, ya que la disponibilidad a pagar es asociable al progreso económico.

3. Aumento de la seguridad vial

Valorar económicamente bienes tan intangibles como la vida humana o daños físicos y morales evitados es complejo y ambiguo, pero resulta imprescindible para la realización de ACBs concluyentes. En la literatura hay dos tendencias principales a la hora de plantear este tipo de valoraciones:

— Cuantificación de los costes asociados. Se valoran los heridos y los siniestros evitados en función de los costos asociados: costes hospitalarios, costes de rehabilitación no médica, pérdida de producción, etc. En esta categoría, se incluirían aquellos métodos que utilizan la cuantía de las indemnizaciones de los seguros para fijar un valor monetario para cada siniestro o herido evitado[24].

— Disponibilidad a pagar. En vez de tomar como referencia los costes asociados que se evitan, nos fijamos en

[21] La encuesta de movilidad cotidiana es la encuesta de movilidad más amplia de Cataluña, y analiza la movilidad de los individuos del área metropolitana de Barcelona (4.345.435 habitantes) en el año 2000.

[22] El porcentaje de tiempo ahorrado asociado a la vuelta a casa es valorado monetariamente como movilidad no obligada.

[23] Algunos análisis coste-beneficio españoles para la construcción de infraestructuras viarias (pe. Romero-Hernández 1999) utilizan valores calculados por el Ministerio de Fomento aunque no son tan actuales y son menos específicos desde el punto de vista territorial.

[24] Este es el método utilizado por Riera (1993) en el análisis coste-beneficio de las rondas de Barcelona.

Cuadro n.º 2

Valoraciones monetarias de los accidentados

	Unidad monetaria	Accidentados mortales	Accidentados graves	Accidentados leves
MOPTMA (1992)	Pesetas (1992)	30.000.000	4.000.000 (1)	
Department of Transport (UK) (2003)	Libras (2000)	1.145.000	128.650	9.920
Riera (1993)	Pesetas (1989)	12.882.409	1.314.000	85.000

Nota: (1) Accidentes no mortales
Fuente: MOPTMA (1992)
Treasury HM (2003)
Riera, P (1993)

el dinero que estarían dispuestos a pagar los individuos con el fin de no sufrir una de estas circunstancias.

En general, hay un amplio consenso entre economistas que los valores monetarios de la seguridad que se utilizan en un ACB deben basarse en las propias preferencias de los afectados: el valor que dan los mismos conductores en evitar un accidente de tráfico. En conclusión, tal y como apunta G. Loomes, tenemos que basarnos en medidas de disponibilidad a pagar para evitar un accidente. A título orientativo, presentamos en el cuadro n.º 2 los diferentes valores que pueden servir como referencia.

El *Department of Transport* británico es el que ha apostado de forma más clara por un análisis contingente: basándose en estudios a partir de entrevistas para revelar de la forma más fiable posible el valor que estarían dispuestos a pagar para evitar una de estas eventualidades (*willingness to pay*). Aunque deba ajustarse su valor por las diferencias de poder adquisitivo, nos parece la fuente más fiable para capturar el aumento de utilidad que genera sobre los individuos la reducción de estos riesgos.

Tomando los valores del *Department of Transport (UK)*, ajustados según el poder de compra en Cataluña, calculamos la reducción de costos asociada al aumento de la seguridad vial, proyectando también estos valores de acuerdo con el aumento esperado de la productividad anual.

6. Impacto ambiental

Una vez calculados los incrementos en las emisiones de CO2, debemos asignar un valor monetario al coste social que comporta el efecto contaminante de las emisiones. Presentamos el referente a nivel estatal (Ministerio de Fomento) y lo contrastamos con el que se utiliza en la Gran Bretaña para realizar evaluaciones económicas (cuadro n.º 3):

Cuadro n.º 3

Valoración monetaria de las emisiones de CO2

	Unidad monetaria	Coste económico por tonelada
Ministerio de Fomento (1996)	Pesetas (1996)	984
Department of Enviroment (UK) (2002)	Dólares (2000)	9,2 (2001-2010) 11,6 (2011-2020)

Fuente: Clarkson R. y Deyes K. (2002). Ministerio de Fomento (1996).

Utilizaremos los valores estimados para Nordhaus y recogidos en el estudio del *Department of Enviroment (UK)*, que una vez ajustados por la diferencia de poder de compra con EE.UU. y actualizados en valores del 2003, resultan ser muy similares.

6. OBTENCIÓN DE LA RENTABILIDAD SOCIAL ACTUALIZADA

La necesidad de aplicar una tasa de descuento parte de la preferencia mayoritaria a consumir en el momento presente respecto a un momento futuro. Aunque en la mayoría de aproximaciones ACB ésta se sitúa alrededor del 6 o el 7%, no hay un acuerdo teórico amplio sobre el tema.

Para fijar la tasa de descuento nos basamos en las estimaciones y planteamientos de Nieves Souto (2003). En la literatura económica existen dos metodologías diferenciadas para calcular la tasa social de descuento:

— El cálculo de la tasa social de preferencia temporal. Se considera que la tasa de descuento social adecuada para valorar los proyectos públicos es la tasa social de preferencia temporal, un indicador que pretende capturar cómo disminuye el valor del consumo para el conjunto de la sociedad con el paso del tiempo.

— El cálculo del coste de oportunidad social de los fondos públicos. Según este método, en cambio, la tasa de descuento social se debería calcular a partir de la estimación del coste de oportunidad social de los fondos públicos: la rentabilidad que se podría obtener de los fondos públicos necesarios para un proyecto en una mejor inversión alternativa. La imperfección de los mercados provoca que este indicador se ajuste exactamente al tipo de interés del mercado.

Utilizar la opción del coste de oportunidad es muy discutible, ya que supondría asumir que la inversión pública desplaza

la misma cantidad de inversión privada que, posteriormente, se destina íntegramente a consumo. Utilizar la tasa social de preferencia comporta unos supuestos menos restrictivos: que la inversión de recursos en proyectos públicos no distorsiona la distribución de recursos entre ahorro y consumo. La metodología sugerida por Nieves Souto (2003) es utilizar la tasa social de preferencia temporal (TSPT). La mejor estimación de TSPT para el caso español se sitúa alrededor del 5%, que es la que utilizaremos en este estudio.

En el cuadro n.º 4 presentamos la cuantificación económica entre el 2003 y el 2020 de cada uno de los beneficios potenciales de suprimir el peaje. En el cuadro n.º 5 se presenta la cuantificación de los costes potenciales. Los resultados globales se presentan en el cuadro n.º 6: los beneficios totales son la suma de los beneficios contemplados en el cuadro n.º 4 y los costes totales son la suma de todos los costes contemplados en el cuadro n.º 5. A partir de los datos del cuadro n.º 6, aplicando la tasa social de preferencia temporal al resultado neto (beneficio-coste) en cada uno de los años, se calcula el valor actual neto de suprimir el sistema de peaje directo.

Cuadro n.º 4

Beneficios económicos derivados de la supresión del peaje

(Euros 2003)

Año	Reducción coste monetario	Valor del tiempo ahorrado	Aumento de recaudación en el imp. especial de hidrocarburos	Aumento de recaudación en IVA sobre hidrocarburos	Aumento de seguridad vial
2003	392.518.025	693.334.746	11.942.442	7.289.125	80.706.945
2004	399.735.009	698.813.050	16.555.477	10.104.712	110.038.995
2005	407.098.385	687.221.547	21.416.200	13.071.477	139.920.388
2006	414.611.481	678.409.148	21.948.502	13.396.370	142.815.116
2007	422.277.712	669.709.752	22.494.337	13.729.523	145.735.878
2008	430.100.577	661.121.911	23.054.063	14.071.154	148.681.775
2009	438.083.664	652.644.193	23.628.043	14.421.485	151.651.952
2010	446.230.654	644.275.188	24.216.652	14.780.746	154.645.619
2011	454.545.321	636.013.500	24.820.276	15.149.171	157.662.079
2012	463.031.536	627.857.753	25.439.310	15.527.001	160.700.747
2013	471.693.267	619.806.590	26.074.161	15.914.485	163.761.186
2014	480.534.588	611.858.668	26.725.247	16.311.878	166.843.139
2015	489.559.674	604.012.665	27.392.996	16.719.442	169.946.567
2016	498.772.810	596.267.272	28.077.851	17.137.446	173.071.692
2017	508.178.390	588.621.201	28.780.263	17.566.167	176.219.046
2018	517.780.922	581.073.176	29.500.700	18.005.889	179.389.529
2019	527.585.033	573.621.942	30.239.639	18.456.904	182.584.465
2020	537.595.468	566.266.257	30.997.573	18.919.513	185.805.673

Fuente: Estimación propia.

Sumario

Cuadro n.º 5

Costes económicos derivados de la supresión del peaje

(Euros 2003)

Año	Desaparición IVA sobre peajes	Desaparición imp. sociedades sobre empresas concesionarias	Compensación a las empresas concesionarias*	Impacto ambiental
2003	33.891.837	71.810.161	313.623.310	1.553.315
2004	34.595.439	73.130.490	318.108.137	2.153.318
2005	35.315.424	74.477.601	322.678.788	2.785.537
2006	36.052.215	75.852.102	327.018.155	2.854.771
2007	36.806.248	77.254.619	331.439.257	2.925.766
2008	37.577.968	78.685.792	335.943.880	2.998.568
2009	38.367.836	80.146.278	340.533.852	3.073.224
2010	39.176.322	81.636.748	345.211.047	3.149.782
2011	40.003.912	83.157.895	349.977.387	5.442.233
2012	40.851.105	84.710.426	354.834.841	5.577.966
2013	41.718.413	86.295.068	359.785.428	5.717.167
2014	42.606.364	87.912.565	364.831.218	5.859.928
2015	43.515.500	89.563.681	369.974.330	6.006.342
2016	44.446.378	91.249.200	375.216.941	6.156.507
2017	45.399.571	92.969.927	380.561.281	6.310.522
2018	46.375.667	94.726.686	386.009.636	6.468.489
2019	47.375.274	96.520.322	391.564.350	6.630.513
2020	48.399.012	98.351.706	397.227.829	6.796.702

Nota: (*) en la hipótesis máxima de compensación (100% de los beneficios esperados)
Fuente: Estimación propia.

Lo primero que constatamos es un volumen de beneficios mayor que el de los costes en todos los años, manteniendo un saldo positivo entre beneficios y costes de aproximadamente 800 millones de euros (del 2003) anuales, muy dependientes éstos de las estimaciones del valor del tiempo ahorrado. Aplicando la tasa social de descuento del 5%, el resultado es un valor actual neto claramente positivo: 9.904.102.846 euros del año 2003.

7. ANÁLISIS DE SENSIBILIDAD DE LAS VARIABLES

En este apartado, con el objetivo de concretar los efectos de las distintas hipótesis utilizadas en el ACB, se visualiza el impacto que tiene la elección y la cuantificación de ciertas variables sobre el resultado final, para a continuación, componer dos escenarios alternativos mediante la asignación de diferentes valores a las variables con un mayor impacto sobre el resultado final.

Las variables consideradas han sido:

Cuadro n.º 6

Costes y beneficios totales

(Euros 2003)

Año	Beneficios totales	Costes totales
2003	1.185.791.283	420.878.624
2004	1.235.247.243	427.987.384
2005	1.268.727.998	435.257.348
2006	1.271.180.616	441.777.243
2007	1.273.947.203	448.425.891
2008	1.277.029.479	455.206.208
2009	1.280.429.337	462.121.189
2010	1.284.148.859	469.173.899
2011	1.288.190.346	478.581.427
2012	1.292.556.347	485.974.338
2013	1.297.249.690	493.516.077
2014	1.302.273.521	501.210.074
2015	1.307.631.345	509.059.854
2016	1.313.327.071	517.069.027
2017	1.319.365.067	525.241.301
2018	1.325.750.217	533.580.478
2019	1.332.487.984	542.090.459
2020	1.339.584.484	550.775.249

Fuente: Elaboración propia a partir de los cuadros n.º 4 y n.º 5.

— *Tráfico inducido*. Los resultados del escenario básico de este estudio asumen un nivel de inducción de tráfico bastante bajo (un 0,5% del tráfico en la red viaria catalana). El análisis de sensibilidad permite ver cómo varían los resultados principales en función de las diversas hipótesis de tráfico inducido (entre el 0% y el 100% del tráfico actual, valor este último todavía muy lejano de los niveles medios en los países más avanzados de la UE: entre el 200 y el 300% en las autopistas).

— *Porcentaje de compensación del beneficio en las empresas concesionarias*. El escenario básico se ha realizado el cálculo con la compensación del 100% del beneficio. Con el cálculo de sensibilidad se pretende visualizar cómo varía la rentabilidad de las AA.PP. en relación con el porcentaje de beneficio para el cual se compensa a las concesionarias.

— *Tasa de descuento*. El valor actualizado de todos los flujos monetarios entre el 2003 y el 2020 cambia obviamente en función de la tasa de descuento que se utiliza para el cálculo. Variar este valor, vista la periodificación de costes y beneficios que podríamos considerar razona-

ble, llevándolo hasta un 10%, casi no modifica los resultados, que son muy poco sensibles a esta variable.

— *Valor de la seguridad viaria y del tiempo*. El valor económico de estos bienes intangibles es uno de los puntos más controvertibles por la ausencia de criterios claros en la asignación de valores, y por el peso que tienen en la determinación de los resultados finales. En el proceso de monetización la variación de términos de seguridad viaria (diferencia en el número de accidentes, muertes, heridos,...) supuestamente derivados de la supresión del peaje directo es considerablemente alta, ya que se han tomado como referencia los valores calculados por el *Department of Transport británico* (UK) ajustados por el poder de compra de Cataluña. Por lo tanto, el análisis de sensibilidad debe ir orientado preferentemente a introducir valores más bajos (del 0 al 100% de los asumidos en los cálculos iniciales). En lo referente al valor del tiempo, parece sin embargo razonable asumir tanto valores más altos como valores más bajos.

En lo que se refiere a los niveles de accidentalidad, los ratios de accidentalidad se postulan constantes a lo largo del período analizado, lo que implícitamente está asumiendo es que los niveles de accidentalidad en todas las vías se mantendrán constantes aunque la distribución del tráfico sea presumiblemente diferente. Éste es un supuesto muy discutible, si tenemos en cuenta que la eliminación de la mezcla de tráficos de corto y de largo recorrido debería tener un impacto positivo sobre la seguridad vial. Es recomendable, por lo tanto, contemplar escenarios en los cuales se reduzcan los niveles de accidentabilidad actuales. Se representan 2 escenarios extremos a este respecto (el más favorable y el más desfavorable a la supresión) a través de un modelo gráfico de sensibilidad. En el eje horizontal representamos el porcentaje de nuevo tráfico inducido y en el vertical el valor actualizado 2003-2020 (expresado en euros 2003) de los cuatro factores esenciales: el VAN social (computando la totalidad de costes y beneficios estimados) y el VAN de las AA.PP. (considerando únicamente las ganancias y las pérdidas para el erario público). También consideramos dos componentes importantes del VAN social como son el coste/beneficio de la seguridad vial y el generado por el ahorro de tiempo.

El escenario más favorable (gráfico n.º 1) es el que otorga un valor más alto al VAN social y al VAN de las AA.PP. Se introducen las siguientes hipótesis: una tasa de descuento del 5%, una compensación del 50% de los beneficios esperados de las concesionarias por eliminar el sistema de peaje directo, una reducción del 75% de los ratios de accidentalidad (el Libro Blanco de la UE supone la reducción de los ratios en un 50% en el año 2010), la utilización de unos valores para monetizar la seguridad vial iguales al 25% del escenario base, que se acercan así también a los valores utilizados por el Ministerio de Fomento en España, y manteniendo el valor económico del tiempo igual al ya utilizado.

El resultado de aplicar estas hipótesis es un VAN social positivo que aumenta a medida que aumentamos los supuestos de inducción de nuevo tráfico. Este impacto es igualmente positivo sobre el

Gráfico n.º 1

Escenario más favorable a la supresión

tráfico inducido	100%
VAN (Euros)	15.807.066.733
VAN AAPP (Euros)	12.813.793.431

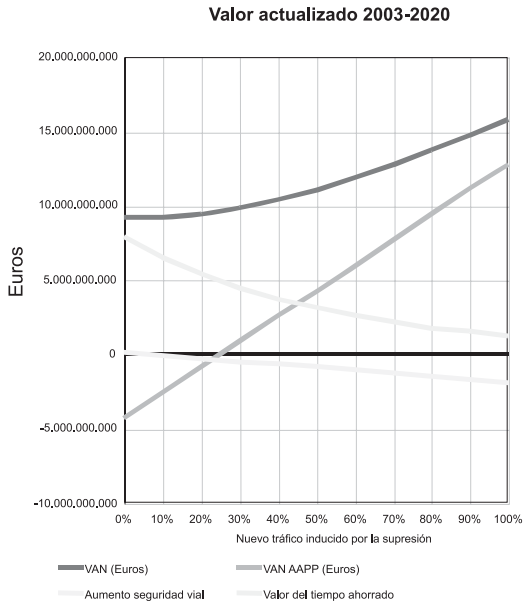
Costes y beneficios anuales (Euros)

Reducción coste monetario	5.524.812.607
Valor del tiempo ahorrado	1.245.231.733
Aumento REC IE carburantes	10.922.747.199
Aumento REC IVA carburantes	6.666.749.406
Aumento seguridad vial	-1.905.888.061
Desaparición IVA peajes	486.002.247
Desaparición IS emp conc	1.010.750.231
Compensación a las concesionarias	3.278.950.696
Impacto ambiental	1.870.882.976

Tasa de descuento	5,0%
% de compensación del beneficio a las concesionarias	50%
Reducción de accidentabilidad 2001-2020	75%
Valor de la seguridad vial	25%
Valor del tiempo	100%

Desglose de la compensación anual a las concesionarias (Euros)

Mantenimiento y conservación de las autopistas	120.619.294
Compensación por lucro cesante	61.544.634
Pago total anual AAPP a las concesionarias	182.163.928



Fuente: Elaboración propia

VAN de las AA.PP., pero con la diferencia de que si asumimos un nivel de inducción de nuevo tráfico inferior al 25% el balance global (2003-2013) para las AA.PP. no es positivo.

Alternativamente, el segundo escenario que presentamos (gráfico n.º 2) es el más desfavorable de todos, presentando tanto una peor rentabilidad social como para las finanzas públicas. Las hipótesis para este escenario son las siguientes: un 5% de tasa de descuento, una compensación del 100% de los beneficios esperados de las concesionarias por eliminar el sistema de peaje directo, una reducción del 25% de los ratios de accidentalidad, la utilización de unos valores para monetizar la seguridad vial y de ahorro del tiempo iguales al escenario base. Observando los valores de los diferentes componentes del gráfico n.º 2 en el eje vertical, se constata que son prácticamente iguales a los del escenario base.

Lo más destacable es que el VAN social de la supresión del sistema de peaje directo es siempre positivo, independientemente de las hipótesis asumidas. El VAN de las AA.PP. (balance presupuestario) es también positivo cuando estimamos un nuevo tráfico inducido de entre el 25 y el

Gráfico n.º 2

Escenario más desfavorable a la supresión

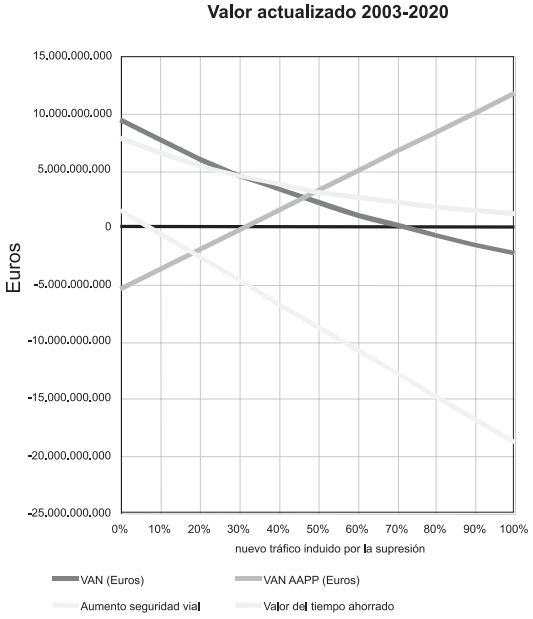
Tráfico inducido	100%
VAN (Euros)	-2.275.181.382
VAN AAPP (Euros)	11.705.990.022

Costes y beneficios anuales (Euros)

Reducción coste monetario	5.524.812.607
Valor del tiempo ahorrado	1.245.231.733
Aumento REC IE carburantes	10.922. 747.199
Aumento REC IVA carburantes	6.666.749.406
Aumento seguridad vial	-18.880.332.768
Desaparición IVA peajes	486.002.247
Desaparición IS emp conc	1.010.750.231
Compensación a las concesionarias	4.386.754.106
Impacto ambiental	1.870.882.976
Tasa de descuento	5,0%
% de Compensación beneficio a las concesionarias	100%
Reducción de accidentabilidad 2020	25%
Valor de la seguridad vial	100%
Valor del tiempo	100%

Desglose de la compensación anual a las concesionarias (Euros)

Mantenimiento y conservación de las autopistas	120.619.294
Compensación por lucro cesante	123.089.268
Pago total anual AAPP a las concesionarias	243.708.561



Fuente: Elaboración propia

30% o superior (en función del escenario). Lo que hace variar este umbral de rentabilidad para las AA.PP. es el porcentaje del beneficio esperado con el que compensamos a las empresas concesionarias que dejan de recaudar el peaje directo (un 50% en el escenario más favorable, y un 100% en el escenario más desfavorable).

8. **CONCLUSIONES**

El análisis demuestra cómo las estimaciones del escenario base indican que la financiación de la alternativa aquí estudiada no es presupuestariamente neutral: cuando no se contempla la posibilidad de tráfico inducido, el flujo de ingresos (recaudación tributaria) no parece ser suficiente para afrontar los gastos. En cambio, con una mayor inducción de tráfico (entre el 25 y el 30% en función del supuesto), el aumento en la recaudación adicional en impuestos sobre hidrocarburos convierte la supresión del peaje en una medida potencialmente asumible para los presupuestos públicos.

En efecto, la sostenibilidad presupuestaria de la supresión del peaje directo, tal y como prevén las actuales restricciones

de déficit y de limitación eventual del nivel de endeudamiento público/PIB (Pacto de Estabilidad) es la clave para que, una vez demostrada su rentabilidad social, poder demostrar que la medida no sólo es plenamente deseable, sino que también es factible. En otras palabras, una vez se demuestra que la supresión del peaje es una medida eficiente (los beneficios derivados son superiores a los costes), su viabilidad dependerá de que pueda ser asumida por parte de las AA.PP.: es decir, que no suponga una carga excesiva sobre sus presupuestos (sostenibilidad presupuestaria). La no inclusión en los cálculos de la recaudación impositiva adicional derivada del aumento potencial de la actividad económica generada por la supresión del peaje directo, convierte por lo demás la sostenibilidad presupuestaria de la medida en una realidad incluso más verosímil.

Sumario

REFERENCIAS BIBLIOGRÁFICAS

ACESA INFRAESTRUCTURAS (2003): *Informe anual 2002.* Acesa Infraestructuras. Barcelona.

ASENSIO J, ROCA O (2001): *Evaluación económica y social del 4º Cinturón de Barcelona.* III Encuentro de Economía Aplicada. Valencia

AUTOPISTAS (2003): *Informe anual 2002.* Autopistas. Barcelona.

AUTOPISTES DE CATALUNYA (2003): *Informe anual 2002.* Autopistes de Catalunya (AuCat). Barcelona

AUTORITAT DEL TRANSPORT METROPOLITÀ (1998): *Els Comptes del transport de viatgers a la RMB l'any 1998.* ATM en colaboración con la UB la UPC. Barcelona.

AUTORITAT METROPOLITANA DEL TRANSPORT (2001): *Enquesta de Mobilitat Quotidiana. 2001.* ATM. Barcelona.

BOARDMAN AE, GREENBERG DH, VINING AR, WEIMER DL (2001): *Cost-Benefit Analysis. Concepts and Practice*. Prentice Hall, New Jersey (EEUU)

CLARKSON R I DEYES K (2002): *Estimating the Social Cost of Carbon Emissions.* Department for Enviroment, Food & Rural Affairs. HM Treasury (UK).

COTO MILLÁN P, INGLADA V (2001): *Costs-benefit analysis of high-speed train: the case of Spain*

de RUS G, INGLADA V (1997): *Cost-benefit analysis of the high-speed train in Spain*. The Annals of Regional Science, 31. Springer-Verlag

de RUS G, ROMERO M, TRUJILLO L (2000): *Financiación privada de carreteras y tarificación óptima ¿Es posible conseguir ambas?* III Encuentro de Economía Aplicada. Valencia

DE RUS G (2001): *Análisis Coste-Beneficio*. Editorial Ariel. Barcelona

ECONOMIC POLICY COMMITTEE (2001): *Budgetary challenges posed by ageing populations* Brussels. 24 october 2001. EPC/ECFIN/655/01-EN final

IDESCAT (2003): *Economia. Preus i salaris* (http://www.idescat.es)

INSTITUT CATALÀ D'ENERGIA (2001): *Campanya d'informació sobre l'etiquetatge energètic i d'emissions de CO2 dels vehicles de turismes nous*. Generalitat de Catalunya

JONES-LEE M, LOOMES G (2001): *The valuation of Health and Safety for Public Sector Decision Making. A Economic Evaluation, from theory to practice*. Springer. Barcelona.

MARTIN F (1997): «Justifying a high-speed rail project: social value vs. regional growth». *The Annals of Regional Science*, 31. Springer-Verlag

MATAS A, RAYMOND JL (2001): *Elasticidad de la demanda de las autopistas de peaje*. III Encuentro de Economía Aplicada. Valencia.

MINISTERIO DE ECONOMÍA (2003): *Precios de carburantes y combustibles. Datos de julio y agosto de 2003.* http://www.mineco.es

MINISTERIO DE FOMENTO (1996): *Manual para la evaluación de inversiones de transporte en las ciudades,* AA.VV. Centro de Publicaciones Mº de Fomento, Madrid.

MINISTERIO DE FOMENTO (2002): *Anuario Estadístico 2001.* http://www.mfom.es

MOPTMA (1992): *Recomendaciones para la Evaluación Económica, Coste Beneficio de Estudios y Proyectos de Carreteras.* Servicio de Planeamiento de Carreteras

MOTJER CERVERO P (1995): *Estudi de la Relació entre Topologia de la Xarxa Viària, Peatge Directe i Mobilitat.* Projecto o tesina de especialitad. Escola Tècnica Superior d'Enginyers de Camins Canals i Ports de Barcelona, UPC.

OILGAS SA (2001): *Enciclopedia Nacional del Petróleo, Petroquímica y Gas, 2001.*

RAYMOND JL I GONZÁLEZ-PÁRAMO M (1988): «Déficit, impuestos y crecimiento del gasto público» *Papeles de Economía Española*, nº37.

RIERA P (1993): *Rentabilidad social de las infrestructuras: Las rondas de Barcelona*. Editorial Civitas, Barcelona.

ROMERO-HERNÁNDEZ MC (1999): «Análisis coste beneficio de un proyecto de inversión en infraestructura de carreteras». *Investigaciones Económicas*, Vol. XXIII.

SEGARRA I BLASCO A. et altri (1999): *Rentabilidad económica y social de la conexión de Tarragona con el ancho de vía europeo: Coste-Beneficio del transporte de mercancías.* Grup d'Investigació Indústria i Territori. Departament de Gestió d'Empreses i Economia, URV.

SERRA DE LA FIGUERA d (1999): *Experiències prèvies de rescat a Europa i Viabilitat d'Aquest a Catalunya*. Mimeo. Institut d'Estudis Territorials. Universitat Pompeu Fabra.

SERRATOSA A et altri (1999): *Les autopistes de peatge a Catalunya. Crònica d'una discriminació inacceptable.* Columna assaig, Barcelona.

Sumario

SOUTO NIEVES G (2003): «Tasas de descuento para la evaluación de inversiones públicas: estimaciones para España». *Papeles de Trabajo nº8/03. Instituto de Estudios Fiscales*. (http://www.minhac.es/ief/principal.htm)

TABASA (2003): *Comptes anuals i informe de gestió de l'any 2002*. Túnels i Accessos de Barcelona Societat Anònima Concessionària (Tabasa). Barcelona.

TREASURY HM (2003): *Appraisal and Evaluation in Central Government (The Green Book)*. London, The Stationery Office. http://www.hm-treasury.gov.uk/economic_data_and_tools/greenbook/data_greenbook_index.cfm

TÚNEL DEL CADÍ S.A.C. (2003): *Comptes anuals i informe de gestió de l'any 2002*. Túnel del Cadí S.A.C. BARCELONA.

Sumario

Estimación del valor monetario de los años de vida ajustados por calidad: estimaciones preliminares

El presente trabajo pretende obtener una estimación del valor monetario del año de vida ajustado por la calidad (AVAC). Para realizar dicha estimación de un modo directo, a través de la disposición a pagar, se puede partir tanto de ganancias pequeñas de salud como de grandes ganancias. Se ha obtenido un intervalo para el valor monetario del AVAC comprendido entre 9.000 y 38.000 €, si bien resultan necesarias investigaciones adicionales sobre los efectos que para la determinación del mismo tienen los estados de salud empleados, su duración y, especialmente, los llamados «efectos de orden», causados por las diferentes secuencias a que dan lugar las combinaciones de ambos.

Lan honen bidez, kalitateak doitutako bizitza-urtearen diru-balioa zenbatetsi nahi da. Zenbatespen hori zuzenean egin ahal izateko, ordaindu beharreko xedapenaren bidez, osasuneko irabazi txikiak edo irabazi handiak har daitezke abiaburutzat. Kalitateak doitutako bizitza-urtearen diru-balioan lortu den tartea 9.000 eta 38.000 euro bitartekoa da. Alabaina, ikerketa gehigarriak egin behar dira enplegatuen osasun-egoerak horren zehaztapena finkatzeko dituen ondorioei buruz, iraupenari buruz, eta, bereziki, «ordena-aldeko eraginei» buruz, hau da, bien nahasketak eragiten dituen sekuentzien eraginez sortutako eraginei buruz.

This paper tries to estimate the monetary value of a quality adjusted life year (QALY). To perform this estimation by using willingness to pay, it is possible to use small health gains and large ones. We have obtained the monetary value of a QALY ranges from ¤9,000 to ¤38,000. However, further research is required in order to test the impact that the health status, duration, and, critically, order-effects, yield on the attempt of estimating a unique monetary value of a QALY.

José Luis Pinto Prades*
Universitat Pompeu Fabra
Jorge-Eduardo Martínez Pérez
Universidad de Murcia

ÍNDICE

1. Introducción
2. Métodos para el cálculo del valor monetario del AVAC mediante la utilización de la disposición a pagar
3. Estudio empírico
4. Resultados
5. Discusión
Referencias bibliográficas

Palabras clave: AVAC; disposición a pagar; efectos de orden

N.º de clasificación JEL: H51, I18, C93

1. **INTRODUCCIÓN**

El paradigma de evaluación en Economía de la Salud es el Análisis Coste-Utilidad, en el cual, los beneficios derivados de cualquier intervención sanitaria se miden a través de los años de vida ajustados por calidad (AVAC). Los AVAC son índices que combinan cantidad y calidad de vida. Así, cada año de vida de una persona se «pondera» o ajusta por la utilidad o bienestar generado por la calidad disfrutada durante ese año de vida. Esto significa que, por ejemplo, 70 años de vida equivaldrían a un número menor de AVAC ya que estamos incorporando a la longevidad los problemas de salud que puede sufrir una persona a lo largo de su vida. Una vez estimados los AVAC asociados a un tratamiento médico o a un programa sanitario, el Análisis Coste-Utilidad interpreta que un tratamiento es tanto más efectivo cuantos más AVAC permite ganar a los pacientes. El hecho de que los recursos existentes sean limitados, lleva aparejada la necesidad de priorizar, de evaluar la efectividad de los tratamientos en relación a su coste. En este contexto, resulta prioritario establecer una estimación de un umbral Coste/AVAC a partir del cual pueda determinarse si una

* Agradecimientos: este estudio ha sido financiado a través del proyecto 081/27/02 de la Agència d'Avaluació de Tecnologia i Recerca Mèdiques de Catalunya. Asimismo, se reconoce el apoyo del Ministerio de Educación a través del proyecto SEJ2004-05079.

determinada intervención de salud es efectiva o no. Para ello resulta necesaria la estimación monetaria del AVAC.

La estimación del valor monetario de la salud, en este caso medida mediante AVAC, puede resultar una herramienta muy útil para la toma de decisiones en aquellos proyectos públicos que tengan como resultado la mejora en la salud de la población. Por ejemplo, puede utilizarse para evaluar políticas medioambientales, ya que muchas de estas medidas tienen como objetivo la reducción en el riesgo de contraer ciertas enfermedades. También puede utilizarse para evaluar políticas de transporte, ya que muchas de ellas tienen como objetivo la reducción en la mortalidad y morbilidad a consecuencia de accidentes. Todas estas políticas requieren la estimación de umbrales que sirvan para decidir cuándo una política que reduzca el riesgo de pérdidas de salud tiene un coste excesivo o tiene un coste razonable.

Pero, el ámbito en el que el valor monetario del AVAC tiene una aplicación más directa es en sanidad. Más en concreto, en la evaluación de tecnologías sanitarias y, especialmente, en la regulación de medicamentos. La estimación del valor monetario del AVAC puede resultar muy útil para decidir cuando una nueva tecnología permite «comprar» salud a un precio razonable. Asimismo, el valor monetario del AVAC puede ayudar a regular el precio de los medicamentos de forma objetiva y (más o menos) predecible. La regla es muy sencilla: no se debe pagar por un medicamento un precio tal que implique «comprar» un AVAC por más de su valor monetario. Tiene que haber una cierta proporción entre el precio marginal de un medicamento y el coste de «comprar» un AVAC mediante dicho medicamento. De esta forma, el precio de un medicamento estaría relacionado con su efectividad y no con su coste de producción. Esto permitiría remunerar de forma adecuada los buenos medicamentos y desincentivar la producción de aquellos que tiene una contribución terapéutica escasa respecto a medicamentos ya presentes en el mercado.

En la actualidad, existen pocos estudios que hayan intentado obtener el valor monetario de un AVAC, ya sea de forma directa (a través de la disposición a pagar) o indirecta (a partir del valor de una vida estadística).

Algunos estudios que han intentado obtener el valor monetario del AVAC lo han hecho de forma indirecta (por ejemplo, Hirth et al., 2000). El concepto de valor de una vida estadística tiene una larga tradición en economía (Jones-Lee, 1969; Viscusi, 1993; Schwab y Soguel, 1995). Para ilustrar este concepto, supongamos que un grupo de 100.000 personas se someten a un tratamiento preventivo que reduce la probabilidad de muerte (o de estar enfermo) en un 1 entre 100.000. El número esperado de muertes (o de personas con problemas de salud) se reduce en 1 unidad dentro del grupo. Supongamos que los individuos que pertenecen a dicho grupo están dispuestos a pagar, por término medio, *m* € por dicha reducción en el riesgo. La disposición a pagar total es, por tanto, *m* € x 100.000. A esta cantidad se la suele denominar disposición a pagar agregada por prevenir una muerte estadística o, más sucintamente, el valor de una vida estadística (VVE). Es importante hacer constar que como el VVE es igual a la disposición a pagar media dividida por la reducción de riesgo,

Sumario

dicho valor puede verse como la media de las relaciones marginales de sustitución entre renta y riesgo. El valor monetario de la vida se ha estimado (Jones-Lee et al, 1985) en 1,5 millones de euros en el contexto de los accidentes de trafico.

Bajo ciertos supuestos restrictivos, el valor monetario de los AVACs puede obtenerse a partir del VVE. El VVE podría ser considerado como el valor presente descontado de los AVAC que un individuo espera tener durante su toda vida. Esto permite obtener el valor de un AVAC a partir del VVE si se utiliza: a) una tasa de descuento apropiada, b) el número de años que se espera vivir y c) la calidad de vida que se espera disfrutar durante estos años. Sin embargo, las estimaciones del VVE se obtienen en contextos relacionados con accidentes mortales donde otros factores, aparte de la salud de las personas, influyen en dicha valoración por ejemplo, la disposición a pagar por evitar un final traumático, lo cual no es un factor que afecte al número de AVAC (Ives et al 1993).

Si exceptuamos estas dificultades, podemos estudiar la metodología que propone un documento publicado por el Ministerio de Sanidad británico (DoH, 1999). Dicho documento se basaba en estimaciones del valor de una vida estadística del Ministerio de Transporte y que rondaban 1,1 millones de euros (en moneda de 1994). Dichas estimaciones estaban realizadas pensando en personas de unos 40 años de edad media, con una esperanza de vida de 76 años. Por tanto, se considera que una persona que muere por accidente de tráfico a los 40 años perdería unos 36 años de vida. El documento supone que una persona de 40 años viviría 26 de los 36 años restantes en buena salud y 10 años en un estado valorado en 0,5 (sobre una escala 0-1), con lo que en realidad perdería 31 AVAC. El valor monetario de un AVAC sería entonces unos 35.000 euros, si no utilizamos ninguna tasa de descuento. Por tanto, puede obtenerse una estimación del valor monetario de un cambio en la salud de la población producido por una política gubernamental, multiplicando el número de AVACs ganados por el valor monetario de un AVAC. Esta cantidad se tendría que comparar con el coste de la política para obtener el beneficio producido por la misma.

En una línea similar al Ministerio de Sanidad, el *Health and Safety Executive* británico (Davies y Teasdale, 1994) calcula el valor monetario de un año en mala salud multiplicando el número de AVAC perdidos y el valor monetario de un año en buena salud. Utilizan la cifra (valores de 1990) de 830.000 € para las pérdidas de salud ocasionadas por una muerte producida en accidente de carretera. Al hacer los cálculos supusieron que un accidente de carretera implicaba la pérdida de 39 AVACs. Con una tasa de descuento del 4%, les llevó a estimar un valor de cada AVAC en 40.000 €.

El mismo método es empleado por la *Federal and Drug Administration (FDA)* para realizar una estimación del valor monetario de los años de vida ganados por los transplantes, así como del valor de los años de vida ganados por diversas regulaciones sobre el tabaco.«Ha habido pocos intentos de estimar un valor monetario de los años de vida ganados. La OMB (*Office of Management and Budget*) sugiere diversas metodologías, incluyendo la anualización con una tasa de descuento adecuada, del valor de la vida estadís-

tica respecto del total de años que constituye la esperanza de vida. Por ejemplo, con una tasa de descuento del 3%, una vida estadística con un valor de $2,5 millones, para una persona con una esperanza de vida de 35 años, un año de vida tiene un valor de $116.500» (Federal Register, 1996, p. 44576).

Usando este valor del año de vida, la FDA estima el beneficio de nuevas disposiciones legales que aumentarían las donaciones de órganos de la siguiente manera: la esperanza estadística de aumento de longevidad y/o de supervivencia del órgano transplantado es de alrededor de 12 años y no los 40 años que subyacen a la mayor parte de las estimaciones del valor de la vida estadística. Usando el concepto más conservador de «año de vida estadístico» ganado, el beneficio de 1.691 transplantes no renales se aproxima a 20.292 años de vida ganados... Aplicando un valor de $116.000 a los años de vida estadísticos ganados mediante transplantes no renales es, aproximadamente de $2.353.872.000« (Federal Register, 1998, p. 33873). Comparando dichos beneficios con los costes de la nueva regulación, la FDA llega a la conclusión de que las nuevas regulaciones sobre transplantes son muy beneficiosas desde el punto de vista de la sociedad.

Esta forma de calcular el valor monetario de los AVAC tiene varias limitaciones que hacen dudar de la validez de las estimaciones así obtenidas. En efecto, un problema grave de dicha metodología es la relación a que da lugar entre el valor monetario de la esperanza de vida y la edad, esto es, el VVE disminuye con la edad. Sin embargo, no está claro que esto refleje las preferencias de la sociedad. De hecho, la literatura (Rodríguez y Pinto, 2002) muestra que si se pregunta a la población en qué periodos de su vida es más importante estar sano, sus respuestas sugieren que en ciertos periodos de la madurez (los que coinciden con la época de atender a los hijos) o incluso de la vejez (después de retirarse del mercado laboral) es especialmente importante estar sanos. Esto no sugiere que el VVE decrece con la edad. Sin embargo, la forma de calcular el valor de los años de vida ganados, como hemos visto, sugiere que entre los 25 y los 50 años, el valor de prolongar la vida disminuye alrededor de un 30%, cuando parece que las preferencias sociales consideran que, al menos, dicha ponderación debería permanecer constante o tener el signo opuesto.

Por otra parte, esta metodología lleva implícito el supuesto de que la disposición a pagar por reducir un riesgo de morir prematuramente se puede equiparar a la disposición a pagar por preservar un número futuro de años, adecuadamente ponderados, en buena salud. Sin embargo, resulta plausible esperar que gran parte de los sujetos cuando manifiestan su disposición a pagar por evitar esa muerte prematura, y en muchos casos traumática, sobre todo en el contexto de los accidentes de tráfico, están valorando a su vez un agregado que recoge tanto el carácter traumático de la muerte, como las valoraciones subjetivas que otorgan los individuos al hecho de ver truncada su existencia y no poder realizar así algunas de las más naturales aspiraciones de cualquier ser humano. Además, gran parte de las estimaciones que se han realizado del VVE se han hecho en ámbitos distintos del estrictamente sanitario y con ello se pueden introducir sesgos adicionales.

Sumario

Debido a estas limitaciones, se ha intentado estimar el valor del AVAC de forma directa, a través de la disposición a pagar. Hay estudios que sin estar diseñados para estimar el valor monetario del AVAC, estiman datos con los que se puede aproximar dicho valor. Así, Cunningham y Hunt (2000), estimaron el valor monetario del AVAC a partir de una serie de encuestas realizadas a 40 pacientes a punto de someterse a un tratamiento ortognático. En estas encuestas se les preguntaba a los pacientes por su disposición al pago por dicho tratamiento, tras haber valorado su propio estado de salud a través de una lotería estándar (*standard gamble*). A continuación, suponiendo un estado de salud perfecta, y una esperanza de vida posterior de 50 años, se obtuvo un valor medio para un AVAC de 506 libras esterlinas (347 €).

Blumenschein y Johannesson (1998), a través de encuestas realizadas a 69 pacientes con asma, y tras valorar su propio estado de salud mediante diversos métodos, obtuvieron la disposición a pagar por un tratamiento que curase dicha enfermedad. Obtuvieron un rango de valores entre los 7.088 dólares y los 45.733 dólares.

Por su parte, Zethraeus (1998) mediante encuestas realizadas a 104 mujeres que seguían un tratamiento hormonal sustitutorio, obtuvo la disposición a pagar por el mismo. Este autor también obtuvo valoraciones del estado de salud de las pacientes con la técnica de la Escala Visual Analógica (*Visual Analog Scale*) y con el método de la Compensación Temporal (*Time Trade-Off*). Bajo el supuesto de salud normal posterior durante treinta años, delimitó un intervalo para el valor monetario de un AVAC de entre 118.400 y 156.100 coronas suecas (aproximadamente 12.900€-16.600€).

Bala et al. (1998), a través de una encuesta realizada con población general con edades comprendidas entre los 65 y 70 años, obtuvo la disposición a pagar por tres tratamientos distintos encaminados a la reducción del dolor derivado del herpes Zoster. Bajo el supuesto de un horizonte vital de 20 años, obtuvieron tres valores distintos para el AVAC: 34.455 dólares; 49.133 dólares y 15.588 dólares.

Por su parte, Olsen y Donaldson (1998), a partir de una muestra de población general de 150 individuos, valoraron tres programas sanitarios con la finalidad de obtener el valor monetario del AVAC. Estos tres programas fueron: un servicio de helicóptero-ambulancia, un desarrollo en la provisión de las operaciones cardiacas, y un mayor desarrollo de las operaciones de cadera. Desde esas valoraciones obtuvieron un valor del AVAC, que variaba entre las 0,2 coronas noruegas y las 6,7 coronas (aprox. 5 céntimos y 1 euro). Dicho rango se explicaba por la utilización o no de una tasa de descuento del 10%.

Hasta el momento, los estudios diseñados específicamente para obtener el valor monetario del AVAC han sido el de Gyrd-Hansen (2003), Byrne et al (2005) y King et al (2005). El estudio de Gyrd-Hansen (2003) utiliza el Análisis Conjunto y obtiene un valor del AVAC de unos $10.000, que es muy inferior al umbral que se ha manejado hasta hoy (entre $50.000 y $100.000 por AVAC). Byrne et al (2005) obtiene un valor monetario del AVAC entre $1.200 y $5.700, mientras que King et al (2005) obtienen un valor del AVAC que oscila entre 12.500$ y 32.000$; estimaciones todas inferiores a cantidades que se venían utilizando, sin mayor fundamento en la literatura (p.ej. $50.000 por AVAC, 30.000£ por AVAC).

Sumario

Otro estudio reciente es el de Lachaine et al (2003) en el que se ajusta una regresión entre la disposición a pagar y mejoras en calidad de vida, pero con una explicación insuficiente.

Finalmente, y para el caso de España, Pinto y Rodríguez (2001), utilizando un enfoque relativamente distinto, intentaron obtener el valor monetario del AVAC a través de encuestas administradas a una muestra de 110 individuos extraídos de la población general. Estos autores obtuvieron la valoración de la disposición a pagar por pequeñas mejoras de salud, calculando el número de AVAC que dichas mejoras de salud suponían, y obteniendo desde ahí el valor monetario de un AVAC. En el estudio de Pinto y Rodríguez se obtuvo un valor monetario que estaba alrededor de los 20.000 euros por AVAC aunque la muestra fue demasiado pequeña (sólo 110 individuos) y poco representativa; por tanto, puede considerarse simplemente como un estudio piloto. El estudio que aquí presentamos puede considerarse una continuación del anterior, intentando superar algunas de sus limitaciones.

2. MÉTODOS PARA EL CÁLCULO DEL VALOR MONETARIO DEL AVAC MEDIANTE LA UTILIZACIÓN DE LA DISPOSICIÓN A PAGAR

Podemos dividir los métodos para estimar el valor monetario del AVAC en dos grandes grupos:

1. Cálculo del valor monetario a través de ganancias pequeñas de salud.

2. Cálculo del valor monetario a través de ganancias grandes de salud.

2.1. Valor monetario del AVAC a través de ganancias pequeñas de salud

Este método parte del supuesto, tradicional en la literatura sobre AVAC, de que el beneficio de un tratamiento sanitario se basa en el número agregado de AVAC ganados, sin considerar su distribución. Por tanto, se estima la disposición a pagar por una cierta ganancia («pequeña») en la salud y se agrega la disposición a pagar necesaria para ganar un AVAC mediante múltiples ganancias pequeñas. El método, por tanto, implica los siguientes pasos:

1. Estimación del número de AVAC ganados mediante la mejora en la salud de una persona. Para ello, se obtendrá la utilidad de un cierto estado de salud [U(Z)] para dicha persona. La mejora en la salud para dicha persona, en términos de AVAC será [1-U(Z)] x t', donde t' es la duración del problema de salud «Z» en términos relativos al año. Obviamente, si la duración del problema de salud es de 1 año, t'=1, los AVAC ganados por no estar un año con el problema de salud Z y estar en buena salud se estiman como [1-U(Z)].

2. Estimación de la disposición a pagar por una mejora segura en la salud [1-U(Z)] x t' a la que denominaremos DP(Z).

3. Estimación del valor monetario del AVAC como:

$$VMAVAC_i = \frac{DP_i(Z)}{[(1 - U_i(Z)) \times t']} \qquad [1]$$

La expresión [1] proporciona el valor monetario del AVAC en el siguiente sentido. Supongamos que un número N de personas tienen una disposición a pagar DP(Z) por una mejora en la salud [1-U(Z)] x t' donde N es tal que:

Sumario

$$N \times [(1-UCZ) \times t'] = 1 \text{ AVAC}$$

Por tanto, N x DP (Z) sería la disposición a pagar total por un AVAC y la expresión [1] nos permite obtener la disposición a pagar por un AVAC. Por ejemplo, si U(Z)=0.9, t'=1 y DP(Z)=500€, N sería 10, ya que 10 x [1-0.9] es 1 AVAC, y la disposición a pagar social por un AVAC sería 10 x 500€. Esto daría un valor monetario del AVAC de 5.000€ que es lo que se obtiene a través de la expresión [1].

La necesidad de que salud [1-U(Z)] x t' tenga un valor pequeño se debe a que, de lo contrario, nos encontraríamos con problemas de límites presupuestarios de los sujetos lo que disminuiría el valor monetario del AVAC (véase Smith, 2005). Por ejemplo, si preguntamos a una persona lo que está dispuesto a pagar por evitar 1 día de migraña al año o 2 días de migraña al año, podríamos esperar que por no estar 2 días al año con migraña pagara el doble que por no estar 1 día. Esto haría que el valor monetario de evitar 1 día de migraña fuera constante. Sin embargo, si preguntáramos lo que está dispuesta a pagar dicha persona por evitar 180 días de migraña, posiblemente no fuera 180 veces más que la disposición a pagar por 1 día menos de migraña ya que dicha persona puede no disponer de suficiente dinero como para pagar 180 veces más. En este caso, el valor monetario de evitar un día de migraña sería decreciente, aunque el valor que el paciente otorga a dicho problema sea constante.

Si queremos estimar el valor monetario del AVAC mediante ganancias grandes tenemos que evitar el problema de la restricción presupuestaria. Esto se puede hacer a través del uso de probabilidades.

2.2. **Valor monetario del AVAC a través de ganancias grandes de salud.**

Cuando la utilidad del estado de salud U(Z) es muy baja y/o la duración t' muy alta, el problema de salud puede ser lo suficientemente grave como para generar problemas de restricciones presupuestarias. En ese caso, se tiene que acudir a situaciones de incertidumbre para obtener el valor monetario del AVAC sin tener el problema mencionado de los límites presupuestarios. El caso límite se produce cuando U(Z)=0 y t' es la esperanza de vida, esto es, la muerte. No tiene sentido, por tanto, preguntar lo que se está dispuesto a pagar por evitar la muerte segura. Lo mismo pasa con estados de salud muy graves. Tal y como se ha explicado en la introducción, cuando estamos ante situaciones de este estilo, se obtiene el valor de la salud mediante preguntas sobre disposición a pagar por reducciones en el riesgo.

En el caso del AVAC, cuando el estado de salud es grave (aunque no implique la muerte), también puede utilizarse el formato de la disposición a pagar por reducir el riesgo de un problema grave de salud. El método requiere realizar los siguientes pasos:

1. Estimación del número de AVAC ganados mediante la mejora en la salud de una persona. Este paso es idéntico al paso número 1 del método anterior.

2. Estimación de la disposición a pagar por una reducción en el riesgo de tener un problema de la salud con utilidad U(Z) y con duración t' a la que denominaremos DP(Z). La reducción en el riesgo la mediremos como una reducción en la probabilidad de tener el problema de salud men-

cionado. Dicha reducción en la probabilidad la denominaremos p*.

3. Estimación del valor monetario del AVAC como:

$$VMAVAC_i = \frac{DP_i(Z)}{[(1 - U_i(Z)) \times t'] \times p^*} \quad [3]$$

Por ejemplo, supongamos que:

— Existe un riesgo de 5 por 1.000 de tener un problema de salud Z y con el tratamiento se reduce a 1 por 1.000. Entonces p* = 4/1.000.

— Supongamos que tener dicho problema de salud implica estar en un estado de salud Z con utilidad [U(Z)] 0,5 y duración (t') 10 años. Por tanto, el número de AVAC ganados por no tener el problema de salud Z sería de 5.

— La disposición a pagar [DP(Z)] por reducir el riesgo de dicho problema de salud es de 600 €.

— El valor monetario del AVAC sería [(600 x 1.000) / 4 x 5] = 30.000 €.

Esta expresión proporciona el valor monetario en el siguiente sentido. Supongamos que hay 1.000 personas que pueden tener el problema de salud Z y cada una de ellas tiene un riesgo del 5 por 1.000. Por tanto, de las 1.000 habría 5 que acabarían teniendo el problema. Supongamos que todas toman un medicamento que hace que, en lugar de 5, únicamente una persona tuviera el problema de salud Z. Si cada una de las 1.000 pagara 600 € por reducir el riesgo, tendríamos que, entre las 1.000 pagarían 600.000 € por reducir el riesgo y 4 personas menos tendrían el problema de salud. Dado que cada una de ellas ganaría 5 AVAC, tendríamos que entre las 1.000 se pagarían 600.000 € por no perder 20 AVAC. Esto implica un valor del AVAC de 30.000 €.

El uso de probabilidades nos permite, por tanto, obtener el valor del AVAC mediante ganancias grandes de salud.

Ambos métodos de estimar el valor monetario del AVAC deben producir el mismo valor ya que el modelo AVAC implica un valor monetario del AVAC constante, bien sea a través de ganancias pequeñas, bien sea a través de ganancias grandes. Por tanto, si el modelo AVAC describe de forma aproximadamente correcta las preferencias de los sujetos, los dos métodos tendrían que conducir a un valor del AVAC similar. El uso de dos métodos distintos no tiene que interpretarse como el uso de dos métodos con validez parcial, sino como una forma de contrastar la validez concurrente de nuestras estimaciones.

2.3. **Métodos de agregación de preferencias**

Las fórmulas 1 y 3 serán utilizadas para obtener el valor monetario del AVAC. Sin embargo, lo que estas fórmulas no nos dicen es la forma de agregar las preferencias de grupos de población. En principio, cabe utilizar dos métodos diferentes:

1. Método de agregación 1.

Se promedia el valor monetario del AVAC de cada individuo, esto es, $VMAVAC_i$ de cada individuo. El valor social es el promedio de estos valores.

2. Método de agregación 2.

Se promedian los valores $DP_i(Z)$ y $U_i(Z)$ y se obtienen los valores medios para toda la muestra. Las ecuaciones 1 y 3 se

aplican sobre estas medias y no sobre los valores individuales de cada sujeto.

Estos dos métodos no siempre darán el mismo resultado. Veamos esto con un ejemplo que presentamos en los cuadros nº1 y nº2.

En el caso 1, los dos métodos dan el mismo resultado ya que 2.333/(1-0,7666) es igual a 10.000 (excepto por redondeo de decimales).

Sin embargo, supongamos que el individuo 1 tiene unas preferencias muy distintas al resto, por ejemplo, supongamos que la utilidad es muy pequeña pero la disposición a pagar muy alta.

En este caso, el promedio de $VMAVAC_i$ sería 173.333 pero el ratio 3.666/(1-0,79666) sería de unos 18.000 €.

En principio, las dos formas de agregar son correctas y la decisión entre una u otra ha de basarse en consideraciones más bien prácticas, como el posible efecto de valores extremos, que influyen más en un método que en otro. Es por ello que en este trabajo optamos por el segundo método de agregación de preferencias.

Cuadro n.º 1

Métodos de agregación. Caso 1

	Individuo 1	Individuo 2	Individuo 3	Media
U(S)	0,9	0,80	0,6	0,766
DP (S)	1.000	2.000	4.000	2.333
Duración t´	1 año	1 año	1 año	
$VMAVAC_i$	10.000	10.000	10.000	10.000

Fuente: Elaboración propia.

Cuadro n.º 2

Métodos de agregación. Caso 2

	Individuo 1	Individuo 2	Individuo 3	Media
U(S)	0,99	0,80	0,6	0,796
DP (S)	5.000	2.000	4.000	3.666
Duración t´	1 año	1 año	1 año	
$VMAVAC_i$	500.000	10.000	10.000	173.333

Fuente: Elaboración propia.

Sumario

3. **ESTUDIO EMPÍRICO**

3.1. **Diseño**

De la explicación realizada en el apartado anterior se deduce que el valor monetario del AVAC se puede obtener a partir del esquema del cuadro n.º 3.

Esto es, podemos obtener el valor monetario del AVAC variando cada uno de los atributos del cuadro nº. 3 o varios a la vez, siempre que dicha variación implique una variación marginal en la salud. De esta forma, no tendremos problemas derivados del límite marcado por la restricción presupuestaria individual. En este articulo, presentamos los resultados de una de las diversas formas de obtener el valor monetario del AVAC. Nos centraremos en obtener el valor monetario variando la duración del problema de salud, manteniendo constante la calidad de vida y asumiendo certeza (probabilidad del efecto – 100%). Por tanto, los atributos que manejamos son duración y coste para dos estados de salud (calidad de vida) distintos. Dichos estados serán el 21212 y el 22223 del Euroqol[1].

Por tanto, el esquema que utilizaremos será el siguiente:

La preguntas que tienen que contestar los sujetos son cuál es la cantidad máxima C_0, esto es, su máxima disposición a pagar por no estar 4 (caso 1) y 2 meses (caso 2) respectivamente, en el estado 22223. Lo mismo con el estado 21212.

Aplicando la ecuación 1 se obtiene el valor monetario del AVAC. Dado que el valor monetario del AVAC requiere conocer la utilidad de los estados de salud, también se obtendrá de cada sujeto la uti-

Cuadro n.º 3

Esquema general del estudio

	Sin tratamiento	Con tratamiento
Calidad de vida	CV_1	CV_0
Duración	t_1	t_0
Probabilidad	P_1	P_0
Coste	C_1	C_0

Fuente: Elaboración propia.

[1] El Euroqol-5D es una medida genérica de los resultados de salud, autoadministrable, que describe un estado de salud como una combinación de 5 dimensiones: *movilidad, cuidado personal, actividad cotidiana, dolor/malestar, ansiedad/depresión*. A su vez, cada una de estas dimensiones puede tomar 3 niveles: *ningún problema, algunos problemas, problemas fundamentales*. Pueden definirse de esta forma 243 posibles estados (35), a los que se añaden los estados «inconsciencia» y «muerte» hasta alcanzar un total de 245 estados de salud. Por ejemplo, el estado «11111» describe ausencia total de problemas en las cinco dimensiones, siendo así el mejor estado del Euroqol-5D.

Sumario

Cuadro n.º 4

Método de estimación del valor monetario del AVAC a través del estado 22223

	Sin tratamiento	Con tratamiento	
		Caso 1	Caso 2
Calidad de vida	22223	22223	22223
Duración	4 meses	2 meses	0 meses
Probabilidad	100%	100%	100%
Coste	0	C_0	C_1

Fuente: Elaboración propia.

Cuadro n.º 5

Método de estimación del valor monetario del AVAC a través del estado 21212

	Sin tratamiento	Con tratamiento	
		Caso 1	Caso 2
Calidad de vida	21212	21212	21212
Duración	4 meses	2 meses	0 meses
Probabilidad	100%	100%	100%
Coste	0	C_0	C_1

Fuente: Elaboración propia.

lidad de los dos estados de salud del Euroqol que utilizamos en este estudio utilizando el método de la lotería estándar.

Este esquema permitirá contrastar diversas hipótesis:

Hipótesis 1 (H1). El valor monetario del AVAC es independiente de la calidad de vida.

Para ello se obtendrá el valor monetario del AVAC mediante los dos estados de salud mencionados, esto es, el 22223 y el 21212.

Hipótesis 2 (H2). El valor monetario del AVAC es independiente de la duración.

Para comprobar si el valor monetario del AVAC es independiente de la duración, se mantendrá todo lo demás constante y se variara la duración para el mismo estado de salud. Esto es, se comparará el valor monetario del AVAC cuando la mejora de salud evita 4 meses con el problema

Sumario

22223 o 21212, o cuando únicamente evita 2 meses. Por tanto, se obtiene el valor monetario del AVAC mediante las siguientes cuatro mejoras en la salud:

- Mejora 1: (22223,4) → (11111,4), una ganancia de 4 meses.
- Mejora 2: (21212,4) → (11111,4), una ganancia de 4 meses.
- Mejora 3: (22223,4) → [(22223,2) + (11111,2)], una ganancia de 2 meses.
- Mejora 4: (21212,4) → [(21212,2) + (11111,2)], una ganancia de 2 meses.

Las tareas que han de realizar los sujetos encuestados son, por tanto, las siguientes:

a. Se obtiene la utilidad de los estados de salud 21212 y 22223 mediante la lotería estándar para cada individuo.
b. Se obtiene su disposición a pagar por una mejora como la del cuadro nº.4, esto es, evitar el estado 22223 durante 4 meses y durante 2 meses.
c. Se obtiene su disposición a pagar por una mejora como la del cuadro nº. 5, esto es, evitar el estado 21212 durante 4 meses y durante 2 meses.

Aplicando la ecuación [1] se puede obtener el valor monetario del AVAC mediante 4 formas distintas, en función de los 2 estados de salud y de las dos duraciones estudiadas.

Hipótesis 3 (H3). El valor monetario del AVAC es independiente del orden de las preguntas.

El método anterior podría estar afectado por, al menos, uno de los múltiples efectos de formato como son los llamados efectos de orden. Dado que obtendremos la disposición a pagar por cuatro ganancias distintas de salud, existen 24 secuencias distintas que se podrían utilizar para contrastar los efectos de orden. En este trabajo se utilizaron las dos secuencias que, en nuestra opinión, podrían dar lugar a mayores efectos de orden. En la secuencia 1, se empieza por la ganancia mas grande [(4 meses, 22223) → (4 meses, 11111)] y se acaba por la mas pequeña [(2 meses, 21212) → (2 meses, 11111)]. En la secuencia 2, el orden es el inverso.

De esta forma, el diseño de nuestro estudio pasa a ser el siguiente:

3.3. **El Cuestionario y el formato de las preguntas**

El cuestionario se estructuró en cuatro partes para cada uno de los grupos. La primera parte era simplemente introductoria, de presentación de la encuesta y de familiarización de los encuestados con los estados de salud. La única tarea que tenían que hacer en esta primera parte consistía en valorar, en una escala del 1 al 10, la gravedad de los dos estados de salud que utilizarían más adelante y que fueron denominados en la encuesta el estado S (22223) y el estado Q (21212). La segunda parte de la encuesta consistió en preguntas del tipo de la lotería estándar, para obtener las utilidades de los estados de salud S y Q. En estas preguntas se pide al encuestado que diga el riesgo de muerte que asumiría para no estar en los estados S y Q de forma crónica el resto de su vida. La tercera parte de la encuesta consistió en las preguntas para obtener la disposición a pagar por diversas ganancias en la salud. La última parte de la encuesta consistió en la recogida de variables socioeconómicas.

Sumario

Cuadro n.º 6

Estructura de las preguntas en los dos grupos.

Grupo A	Grupo B
1. (22223,4) → (11111,4)	1. (21212,4) → [(21212,2) + (11111,2)]
2. (21212,4) → (11111,4)	2. (22223,4) → [(22223,2) + (11111,2)]
3. (22223,4) → [(22223,2) + (11111,2)]	3. (21212,4) → (11111,4)
4. (21212,4) → [(21212,2) + (11111,2)]	4. (22223,4) → (11111,4)

Fuente: Elaboración propia.

El formato de las preguntas para obtener la disposición a pagar, fue similar en los grupos A y B. Se pedía al sujeto que pensase que tenía unos síntomas que podían hacer que tuviera una cierta enfermedad que le llevaría a estar en la situación S o Q el resto de su vida. Se le decía que existía una medicina (la X) que les curaría si la tomaban durante 1 año, pero que no tenía un efecto inmediato. Se les decía que tardaría unos 4 meses en hacer efecto. Dicha medicina no les suponía ningún coste adicional. Se les decía que existía otra medicina (la A) que era mejor que la X, pero que costaba algo de dinero. Por tanto, se les preguntaba cuanto estaban dispuestos a pagar al mes, durante 1 año, para utilizar la medicina A y no la X.

4. **RESULTADOS**

Se realizaron un total de 160 encuestas (80 personas por grupo) de forma individualizada (entrevistas cara a cara) por entrevistadores a quienes se formó sobre la naturaleza de la encuesta. Las entrevistas se realizaron en la ciudad de Barcelona y en su cinturón durante los meses de marzo, abril y mayo de 2005. En cada uno de los grupos, se aplicó un sistema de cuotas por edad y sexo para que los grupos fueran comparables entre sí.

Los resultados para los grupos A, y B los podemos observar en el cuadro nº. 7.

El cuadro nº. 7 permite obtener las siguientes conclusiones:

1. La disposición a pagar tiene cierta racionalidad en cada uno de los grupos ya que aumenta con el tamaño del beneficio.

2. La disposición a pagar no es proporcional a la duración del beneficio para el grupo A pero si para el grupo B, esto es, para el grupo A la disposición a pagar por 4 meses no es el doble que para dos meses, aunque para el grupo B si que ocurre esto (las diferencias que se observan para el grupo B no son estadísticamente significativas).

3. Existe una gran influencia de los efectos de orden. Cuando se empieza por la ganancia mas pequeña (grupo B) la disposición a pagar de las cuatro mejoras

Cuadro n.º 7

Disposición a pagar (€/mes) y utilidades.

	Grupo A	Grupo B
Disposición a pagar por		
(22223,4) → (11111,4)	157	78
(21212,4) → (11111,4)	112	51
(22223,4) →[(22223,2) + (11111,2)]	102	45
(21212,4) →[(21212,2) + (11111,2)]	69	28
Utilidades		
Estado 22223	0,72	0,68
Estado 21212	0,87	0,83

Fuente: Elaboración propia.

es mas pequeña que cuando se empieza por la ganancia mas grande (grupo A).

A partir de los datos del cuadro n.º 7 puede obtenerse el valor monetario del AVAC para cada uno de los 8 casos anteriores[2]. En el cuadro n.º 8 podemos observar que:

1. El valor monetario del AVAC es similar en las cuatro ganancias utilizadas para el grupo B pero es bastante distinto para el caso del grupo A.

2. El valor monetario del AVAC es siempre mayor para el grupo A que para el grupo B.

3. El valor monetario del AVAC a través del estado de salud 21212 tiende a ser mayor que el estimado a través del estado 22223 en el grupo A pero no en el B.

5. **DISCUSIÓN**

Este estudio ha intentado obtener lo que hemos denominado el valor monetario de un AVAC. Estas estimaciones se han realizado utilizando únicamente uno de los varios métodos posibles por lo que las conclusiones han de ser, obligatoriamente limitadas al método seleccionado. En la actualidad se esta trabajando en otros métodos que podrán aportar mayor información sobre el valor monetario del AVAC. Teniendo en cuenta esta limitación, los resultados obtenidos permiten obtener unas conclusiones preliminares sobre el valor monetario del AVAC.

[2] Así por ejemplo, el valor monetario del AVAC obtenido en el primero de los escenarios considerados en el Grupo A, se obtendría dividiendo el producto de la disposición a pagar media (157 €) por 12 meses, entre la ganancia de utilidad (1-0.72=0.28) multiplicada por su duración respecto a la anualidad (4/12). Se pueden observar pequeñas diferencias derivadas del número de decimales empleados para el cálculo.

Sumario

Cuadro n.º 8

Valor monetario del AVAC (€).

Mejoras de salud	Grupo A	Grupo B
(22223,4) → (11111,4)	20.158	8.882
(21212,4) → (11111,4)	30.574	10.686
(22223,4) → [(22223,2) + (11111,2)]	26.253	10.197
(21212,4) → [(21212,2) + (11111,2)]	38.069	11.640

Fuente: Elaboración propia.

1. El valor monetario del AVAC obtenido oscila entre 9.000 € y 38.000 €. Obviamente, hay dos formas de ver estos resultados. Por un lado, de forma negativa, ya que las estimaciones tienen un rango muy amplio. Por otro, de forma positiva, ya que al menos se disponen de umbrales estimados empíricamente sobre el valor monetario del AVAC. Hasta el momento, se han estado usando estimaciones realizadas con muy poca base empírica.

2. Es preocupante el hecho de que, al menos en un grupo, el valor monetario del AVAC sea muy distinto en función del estado de salud utilizado y en función de la duración del beneficio de salud. Ambos elementos no deberían influir a la hora de obtener un único valor del AVAC. Futuras investigaciones deberán intentar arrojar más luz sobre la existencia de un valor único del AVAC, como parece que indica el grupo B, o sobre la existencia de un valor cambiante en función de la duración o del estado de salud utilizado como indica el grupo A. Si este último fuera el caso, la labor empírica necesaria para asociar valores monetarios a ganancias de salud sería muy alto, ya que harían falta estimaciones distintas para ganancias distintas.

3. El aspecto que en nuestra opinión es más preocupante de todos es la gran influencia de los efectos de orden. Mientras que puede haber razones teóricas, basadas en las preferencias individuales, para no asignar un único valor monetario del AVAC a ganancias distintas (en función de la calidad de vida o la duración) no existe ningún argumento teórico para utilizar un valor monetario del AVAC distinto en función del orden de las preguntas. Si este efecto se repitiera en posteriores estudios, habría que comprobar hasta que punto afectan a las estimaciones del valor monetario del AVAC. Hay, al menos, dos posibilidades:

a. Si las estimaciones varían con el orden dentro de un rango que puede considerarse aceptable, se podrá decir que el valor monetario del AVAC esta dentro de dicho rango. Todas aquellas tecnologías con un coste por AVAC por debajo del rango podrían ser financiadas públicamente y aque-

llas que estén por encima, al contrario. Obviamente, las decisiones de financiación publica de tecnologías médicas están sujetas a otras consideraciones, pero esta sería la información que el análisis coste-efectividad podría proporcionar.

b. Si las estimaciones hacen variar mucho el rango y se considera poco informativo para orientar decisiones de política sanitaria, pensamos que hay dos alternativas posibles:

   1 Realizar estudios cualitativos con los que se intente conocer mejor las razones por las que dichos efectos se producen, para intentar así eliminarlos.

   2 Cambiar los métodos usados para obtener el valor monetario del AVAC. Por ejemplo, métodos como la disposición a pagar obtenida mediante preguntas dicotómicas o métodos basados en Análisis Conjunto pueden ser menos exigentes cognitivamente para los encuestados, lo que puede reducir el problema generado por los efectos de orden.

En resumen, estas primeras estimaciones del valor monetario del AVAC apuntan a un rango que oscila entre los 9.000 € y los 38.000 € por AVAC. Las variaciones ocasionadas por los estados de salud, la duración del efecto y (principalmente) los efectos de orden, muestran la necesidad de realizar más estudios que permitan contrastar la fiabilidad de estas primeras estimaciones preliminares.

Apéndice 1

Estados de salud utilizados en el estudio.
Estado (11111)

— *No tengo* problemas para caminar
— *No tengo* problemas para lavarme o vestirme solo
— *No tengo* problemas para realizar mis actividades cotidianas (p.ej., trabajar, estudiar, hacer las tareas domésticas, actividades familiares o actividades durante el tiempo libre)
— *No tengo* dolor o malestar
— *No estoy* ansioso o deprimido

Estado S (22223)

— *Tengo algunos* problemas para caminar
— *Tengo algunos* problemas para lavarme o vestirme solo
— *Tengo algunos* problemas para realizar mis actividades cotidianas (p.ej., trabajar, estudiar, hacer las tareas domésticas, actividades familiares o actividades durante el tiempo libre)
— *Tengo moderado* dolor o malestar
— *Estoy muy* ansioso o deprimido

Estado Q (21212)

— *Tengo algunos* problemas para caminar
— *No tengo* problemas para lavarme o vestirme solo
— *Tengo algunos* problemas para realizar mis actividades cotidianas (p.ej., trabajar, estudiar, hacer las tareas domésticas, actividades familiares o actividades durante el tiempo libre)
— *No tengo* dolor o malestar
— *Estoy moderadamente* ansioso o deprimido

Sumario

REFERENCIAS BIBLIOGRÁFICAS

BALA M.V., WORD L., ZARKIN G., NORTON E., GAFNI A. Y O´BRIEN, B. (1998): «Valuing outcomes in health care: a comparison of willingness to pay and quality adjusted life years». *Journal of Clinical Epidemiology*, nº 51(8), págs. 667-676.

BLUMENSCHEIN K. y JOHANNESSON, M. (1998): «The relationship between quality of life instrument, health state utilities and willingness to pay in patients with asthma». *Annals of Allergy, Asthma and Immunology*, nº 80(2), págs. 189-194.

BYRNE M., O'MALLETY K. y SUAREZ-ALMANZOR, M. (2005): «Willingness to pay per Quality-Adjusted Life Year in a study of knee osteoarthritis». *Medical Decision Making*, nº 25, págs. 655-666.

CUNNINGHAM S.J. y HUNT N. (2000): «Relationship between utility values and willingness to pay in patients undergoing orthognathic treatment». *Community Dental Health*, nº 17, págs. 92-96.

DAVIES N.V., y TEASDALE P.(1994): *The cost to the British economy of work accidents and work-related ill health*. HMSO .London.

DEPARTMENT OF HEALTH (1999): *Economic appraisal of the health effects of air pollution.* Stationary Office.

FEDERAL REGISTER (1998) vol. 63, nº. 119, págs. 33856-33875.

FEDERAL REGISTER (1996) vol 61, nº. 168, pág. 44576.

GYRD-HANSEN, D. (2003): «Willingness to pay for a QALY». *Health Economics*, nº12 (12), págs. 1049-1060.

HIRTH R.A., CHERNEW M.E., MILLER E., FENDRIK M.F. y WEISSERT W.G. (2000): «Willingness to pay for Quality-Adjusted Lifer Year: In search of a Standard». *Medical Decision Making*, nº 20, págs. 332-342.

IVES D., SOBY B., BALL D., y KEMP R. (1993): «Revaluation of non-fatal casualty costs: a report of the application of the relative utility loss approach (RULA)». TRL Contractor Report 329.

JONES-LEE, M.W.(1969): «Valuation of Reduction in Probability of Death by Road Accident». *Journal of Transport Economics & Policy* , nº 3 (1), págs. 37-47.

JONES-LEE, M.W., HAMMERTON M. y PHILIPS P.R. (1985): «The value of safety: Results of a national sample survey». *Economic Journal*, nº 95, págs. 49-72.

KING J. TSEVAT J., LAVE J.,y ROBERTS M. (2005): «Willingness to pay for a Quality-Adjusted Life Year: Implications for societal health care resource allocation». *Medical Decision Making*, nº 25, págs. 667-677.

LACHAINE J., LAURIER C. y CONTANDRIOPOULOS A.P. (2003): «Defining monetary values for quality of life improvements: an exploratory study». *Pharmacoeconomics,* nº 21(12), págs. 865-74.

OLSEN, J.A. y DONALDSON, C. (1998): «Helicopters, hearts and hips: using willingness to pay to set priorities for public sector health care programmes». *Social Science and Medicine*, nº 46, págs. 1-12.

PINTO PRADES, J.L. y RODRIGUEZ MONGIÓ, R. (2001): «¿Cuánto vale la pena gastarse para ganar un año de vida ajustado por calidad? Un estudio empírico» en PUIG J., DALMAU E. y PINTO J.L., *El valor monetario de la salud.* Springer-Verlag, Barcelona.

RODRÍGUEZ-MÍGUEZ E. y PINTO-PRADES J.L. (2002): «Measuring the social importance of concentration or dispersion of individual health benefits». *Health Economics*, nº 11, págs. 43-53.

SMITH, R. (2005): «Sensitivity to scale in contingent valuation: the importance of the budget constraint». *Journal of Health Economics,* nº 24 (3*),* págs. 515-552.

SCHWAB, N. y SOGUEL,N.(eds.) (1995): *Contingent Valuation, Transport Safety and the Value of Life.* Kluwer Academic Publishers. Boston.

VISCUSI, K. (1993): «The value of risks to life and health». *Journal of Economic Literature*, nº 31 (4), págs. 1912-1946.

ZETHRAEUS, N. (1998): «Willigness to pay for hormone replacement therapy«. *Health Economics*, nº 7, págs. 31-38.

Sumario

Política europea de empleo para mujeres inactivas: ¿cómo sabremos si estamos acertando?

El objetivo del análisis y de la evaluación es proporcionar información objetiva a los gestores de los programas y a los representantes políticos sobre los costes y los efectos de los programas. El caso que se presenta a continuación —en el marco de la iniciativa europea Equal— ilustra esta idea. Se trata de un programa europeo en el que la administración responsable gestiona correctamente y obtiene unos resultados aceptables. Sin embargo, una investigación basada en entrevistas en detalle a treinta usuarias seleccionadas aleatoriamente pone de relieve que la evaluación estandarizada no responde a preguntas fundamentales sobre el valor público del programa. Así pues, se cuestiona si la evaluación que se ha ido desarrollando en Europa en los últimos años ha sido capaz de adaptarse a cada proyecto y si puede estar produciendo información útil para la comparación básica entre proyectos pero no para tomar decisiones estratégicas sobre las políticas.

Azterketaren eta ebaluazioaren helburua du programetako kudeatzaileei eta ordezkari politikoei programen kostu eta ondorioei buruzko informazio objektiboa ematea. Jarraian aurkezten den kasua —Europako Equal ekimenaren esparruan— ideia horren argigarri da. Europa mailako programa honen bidez, administrazio arduradunak kudeaketa-lanak zuzen egin eta onartzeko moduko emaitzak lortzen ditu. Hala eta guztiz ere, hogeita hamar erabiltzaile ausaz aukeratu dira, haiei zehaztasun handiko elkarrizketa egiteko eta elkarrizketon gaineko ikerketak agerian jarri du ebaluazio estandarizatuak ez diela programaren balio publikoari buruzko funtsezko galderei erantzuten. Litekeena da azken urteotan Europan egin den ebaluazioa proiektu bakoitzari egokitzeko gauza ez izatea ere, eta proiektuen arteko oinarrizko konparaketa egiteko informazio baliagarria eman arren, politikei buruzko erabaki estrategikoak hartzeko beharrezko informazioa, berriz, ez ematea.

The aim of policy analysis and evaluation is to provide objective information to program managers, decisions makers and users on the costs and effects of public programmes. The following case —a program developed under the framework of the European initiative EQUAL— illustrates this idea. It is about a European programme in which the responsible administration manages correctly and achieves acceptable results. However, an investigation based on detailed interviews to thirty randomly selected users shows that the usual standardised evaluation does not respond to fundamental questions about the public value of the programme. It can be questioned therefore that, the evaluations which have developed in Europe during the last years is able to adapt to each project and produce useful information. The evidence they collect may be useful for the basic comparison of projects but not to take strategic decisions.

Xavier Ballart Hernández
Universidad Autónoma de Barcelona-Bellaterra

ÍNDICE

1. Introducción
2. La evaluación en el contexto europeo
3. El estudio de caso
4. Conclusiones sobre la evaluación de políticas de empleo europeas
Referencias bibliográficas

Palabras clave: evaluación, Europa, mujeres, inactividad, gestión, servicios públicos

N.º de clasificación JEL: H43, H53, I3, J16

1. **INTRODUCCIÓN**

La gestión de los servicios sociales incluye entre sus componentes principales las ideas de análisis y evaluación. Esta idea se fundamenta en el paradigma clásico de la planificación racional que entiende que es posible aplicar métodos y técnicas de las ciencias sociales a los programas públicos para entender qué pasa como resultado de la intervención pública en la realidad social. El objeto del análisis y de la evaluación sería por tanto proporcionar información objetiva a los gestores de los programas y a los representantes políticos sobre los costes y los efectos de los programas (Weiss, 1998, Wholey, 1983). A menudo también se habla de análisis de procesos —por ejemplo de gestión intergubernamental en el caso europeo— si bien la aplicación de «rigor» en el análisis suele reducirse a recoger e interpretar algunos indicadores objetivos que aportan información aproximada sobre resultados, actividades o recursos empleados (Hatry, 1999) que se espera sean útiles para fundamentar un juicio sobre la eficacia o la eficiencia de la intervención pública.

La práctica europea en política agrícola, desarrollo regional, empleo o formación suscita, sin embargo, algunas dudas sobre la utilidad de los análisis y evaluaciones a las que ha dado lugar. A continuación se analiza la evolución que ha seguido la práctica de la evaluación en el contexto de las políticas europeas con el objeto de identificar la trayectoria que han seguido los Estados miembros, las regiones y los gobiernos locales induci-

dos por las instituciones europeas. Con este artículo se persigue señalar los riesgos de institucionalización de lo que puede ser una nueva forma de evaluación-justificación que, lejos de aportar valor, puede incluso suponer un obstáculo a la finalidad principal de aprender y alimentar un debate abierto y sincero sobre el valor o el acierto de adoptar determinadas políticas.

El caso que se presenta a continuación ilustra esta idea. Se trata de un programa europeo en el que la administración responsable gestiona correctamente y obtiene unos resultados aceptables. Sin embargo, una investigación simple sobre la base de entrevistas en profundidad a treinta usuarias seleccionadas aleatoriamente pone de relieve que la evaluación estandarizada no responde a preguntas fundamentales sobre la el valor público del programa. Podría pues estar ocurriendo que de la misma manera que la administración tradicional tiene dificultades para dar un tratamiento diferenciado a sus usuarios, la evaluación que se ha ido desarrollando en Europa en los últimos años tampoco sea capaz de adaptarse a cada proyecto y esté produciendo información útil para la comparación básica entre proyectos pero no para tomar decisiones estratégicas sobre las políticas.

2. LA EVALUACIÓN EN EL CONTEXTO EUROPEO

1. Breve referencia a la evolución de la evaluación en Europa

Desde los años setenta del siglo pasado, las administraciones públicas han intentado desarrollar modelos económicos formales para evaluar los impactos de las políticas sociales. Sin embargo, la formalización excesiva y la sofisticación metodológica han supuesto un fracaso relativo de estas aproximaciones (Mairate 2003) que se ha traducido en una baja utilización de los resultados de los análisis.

A partir de los años ochenta, se empiezan a desarrollar otros modelos menos sofisticados metodológicamente pero más preocupados por la medida en que la inversión de dinero público se traduce en servicios eficientes para los ciudadanos (Mairate, 2003). Esto supone poner mayor atención en quién se beneficia de los programas y en qué medida los beneficiarios coinciden con el objetivo propuesto y obtienen servicios tangibles que puedan suponer una mejora en su situación. Esta técnica, no tiene en cuenta el contexto en el que se desarrollan las políticas ni cómo se producen los efectos en los beneficiarios por lo que tiene una limitada capacidad explicativa sobre la naturaleza de los problemas y su interacción con la intervención pública.

Más recientemente, a partir de los años noventa, hemos oído hablar de otros modelos posibles en evaluación como por ejemplo la evaluación participativa o la evaluación en base a la teoría del programa (*theory based* en inglés)[1]. Sin embargo, ambos tipos de evaluación son raramente practicados. Si bien es bastante normal la participación pasiva de usuarios o profesionales de un sector en la evaluación porque son invitados a

[1] Por ejemplo en los debates de la European Evaluation Society. Ver la *Newsletter* de la sociedad europea en www.europeanevaluation.org.

responder a un cuestionario o a hablar sobre su experiencia, resulta mucho menos frecuente la presencia activa de *stakeholders* internos y externos en una instancia de evaluación o que se les transfiera el poder de decisión sobre la evaluación de una política (Guljt y Gaventa, 1998) En el mismo sentido, la idea de evaluar en base a la teoría que subyace en una intervención pública tiene mucho sentido (Weiss, 2000) pero su aplicación al caso concreto se complica enormemente ya que obligaría a comprobar como a raíz de una actividad pública se va produciendo una cadena de procesos complejos (difusión de información, cambios de actitudes, cambios de comportamientos) hasta producir el resultado esperado .

2. **La política europea de evaluación**

La política europea se fundamenta en una serie de principios fundamentales que suponen una cierta racionalización de la gestión de los programas y fondos. Algunos de estos principios son la programación plurianual —lo que permite dejar actuar a los gestores de los programas y esperar un tiempo prudencial antes de analizar los resultados—, la financiación compartida —y por tanto la implicación de los actores territoriales en la política europea y en la adopción de aquellas reformas que se estimen necesarias para su éxito—, la concentración de recursos en determinadas zonas —con el propósito de que las regiones menos avanzadas atrapen a las que lo están más— o la formación de redes sobre la base del partenariado público-privado —fórmula sobre la que se apuesta para generar desarrollo económico local y empleo—.[2]

En el ámbito de los fondos estructurales, por ejemplo, una de las principales condiciones impuestas para doblar los presupuestos en 1999 fue que se sistematizara la evaluación de las políticas y se vinculara de una forma más efectiva al ciclo presupuestario. En la llamada Agenda 2000 se introdujo la «*performance reserve*» (el 4% de los fondos aprobados a medio plazo) para mejorar la función de seguimiento y de evaluación de los programas.

La Comisión Europea adoptó un ambicioso programa de institucionalización de la evaluación con el establecimiento de unidades de evaluación en cada una de las Direcciones Generales bajo la coordinación de la Dirección General de Presupuestos. La implantación de la evaluación fue acompañada de programas de formación y de desarrollo de métodos y técnicas[3]. En 2002, se aprobaron directrices generales de aplicación general —«*evaluation standards*»— y se invitó a los estados miembros a desarrollar programas similares en sus administraciones nacionales. En el mismo sentido, algunos Estados miembros siguieron políticas similares de formación, preparación metodológica e institucionalización de la evaluación en su propia administración central o regional (por ejemplo, Italia).

---

[2] La Comisión Europea ha alimentado el debate sobre estos principios encargando estudios sobre su aplicación como por ejemplo en el estudio sobre el principio del partenariado realizado en 1999: «*The Thematic Evaluation of the Partnership Principle*», The Tavistock Institute, 1999.

[3] El programa MEANS en la DG de Política Regional.

La ampliación de las políticas europeas y la implantación de unidades transversales creadas con la misión de evaluar programas produjeron un considerable desarrollo de la actividad de evaluación y, con ella, la aparición de un número importante de informes de evaluación. La demanda de estudios con esta rúbrica se ha multiplicado en los últimos años, lo que ha contribuido a crear un mercado cada vez más amplio al que acuden las empresas de consultoría que han visto en éste un nuevo sector de actividad. La misma Comisión ha participado en la creación de una Sociedad Europea de Evaluación y ha alentado la creación de sociedades nacionales donde suelen encontrarse los profesionales de este nuevo sector.

Resulta difícil evaluar la política de evaluación europea lo que en cualquier caso no es el propósito de este artículo. La principal idea que se pretende explicar es que, si bien el impulso de la función de evaluación puede haber sido útil para la negociación que se produce en el contexto europeo cuando se distribuyen recursos, está bastante menos claro que haya servido para la reflexión estratégica sobre las políticas europeas (la teoría detrás de las intervenciones) o sobre los métodos europeos (como, por ejemplo, el principio del partenariado para producir desarrollo económico local).

Siguiendo con este argumento, la evaluación no trata de medir el valor que aportan los programas públicos a la sociedad o si funcionaron en el caso concreto —lo que suelen pedir con insistencia los representantes políticos a los gestores públicos— sino que más bien sirve para integrar, en un contexto que es percibido como de distribución de recursos, conocimiento especializado sobre la política bajo la forma de expertos sectoriales que firman los informes de evaluación. Esto es lo que permite a los administradores europeos afirmar que la distribución de los recursos se hace en base a un proceso de evaluación.

Este concepto de evaluación-negociación contrasta con la idea de evaluación-aprendizaje y de experimentación en relación con problemas que, por lo general, ni siquiera están bien definidos (Dente, 2003). Es lógico esperar que en los ámbitos que opera la política europea no se sepa muy bien si las políticas aprobadas van a funcionar, por ejemplo, en desarrollo económico rural, en transformación de barrios degradados en ciudades o en activación y empleo de algunos grupos con especiales dificultades. En estos ámbitos no está muy claro que desarrollar determinadas actividades —normalmente definidas de forma muy genérica como por ejemplo, actividades económicas alternativas a la agricultura tradicional, nuevos yacimientos de empleo, lucha contra la feminización de la pobreza— después vaya a traducirse en efectos positivos y duraderos para la sociedad. Si un ciudadano se acerca a un hospital a que le pongan una vacuna, ello va a mejorar su salud pero si se invierte en desarrollo económico no esta claro que se esté vacunando a la sociedad en la misma medida. Incluso si se invierte en servicios tangibles como una carretera, no se sabe con certeza si los efectos positivos de la infraestructura se van a producir o si van a compensar la pérdida ambiental que necesariamente comporta su construcción.

3. **Las dificultades metodológicas**

En el caso de las políticas sociales y de las europeas en particular, las dificul-

tades metodológicas son importantes. Algunos de los principales problemas tienen que ver con el análisis de la causalidad, de la interacción entre intervención y contexto o de la medida en que se producen escurridizos efectos buscados por la política europea como por ejemplo, la «adicionalidad» o la «sostenibilidad». Otros problemas son menos profundos pero generan importantes dificultades a la utilización efectiva de los análisis, como por ejemplo, la falta de adecuación de la temporalidad de la evaluación o la limitada influencia de los análisis en las decisiones presupuestarias. Aún otros problemas se refieren a cuestiones más operativas, en principio más fáciles de resolver, como por ejemplo la falta de objetivos cuantificados, la disponibilidad de datos comparativos o de estándares que permitan establecer comparaciones.

Uno de los principales riesgos es precisamente que todo el esfuerzo y el debate en evaluación se acabe centrando en indicadores de actividad y del número de beneficiarios en perjuicio de otros debates más interesantes como el de las categorías de estas actividades y beneficiarios. Por ejemplo, existen muchos datos estadísticos sobre niveles de ocupación y desempleo pero mayor ambigüedad sobre las categorías de ocupados —niveles de seguridad en el empleo, niveles de retribución, discriminación por sexo— o de desocupados —registrados en servicios de empleo, registrados que además buscan activamente trabajo, registrados que no buscan trabajo por haber desistido tras no encontrar trabajo, etc—. Este mismo problema de categorías se podría aplicar a otros ámbitos donde resulta difícil evaluar porqué no están bien definidas las principales variables.

La repuesta práctica a estos problemas que suelen dar los gestores profesionales refleja una visión «*top-down*» no ajena al modelo burocrático clásico ya que consiste en forzar a todas las partes a trabajar con objetivos claros y convertir la evaluación en un seguimiento regular de las actividades. Desde esta perspectiva se entiende que la evaluación debe ser una «herramienta» de gestión y debe evitar experiencias que acaban siendo insatisfactorias porque se acercan más a la investigación que a la producción de información en tiempo útil para los gestores profesionales que deben tomar decisiones sobre financiación de proyectos que compiten por unos mismos recursos.

Algunas metodologías importadas del control de gestión toman en consideración el hecho de que las políticas tengan objetivos diversos y de distinto alcance, si bien generalmente relacionados, lo que se traduce en un trabajo de exploración de los valores de los distintos actores o *stakeholders* y de atribución de pesos a los principales objetivos estratégicos.[4] Sin embargo, el producto de la evaluación continua siendo una matriz de datos que permita comparar el rendimiento de varios proyectos en relación con los principales productos esperados a pesar de que los datos aporten poca información sobre la medida en que se avanza en relación con objetivos estratégicos relevantes o sobre qué es lo que habría que modificar en la teoría del programa para profundizar en los procesos de cambio social deseados. En otras palabras, las administraciones son capaces de identifi-

[4] En este sentido la consultora Accenture ha desarrollado la metodología PSV.

car los proyectos más productivos en términos de beneficiarios o de actividades para beneficiarios (certificados de formación, prácticas, talleres de orientación profesional) pero no de responder a cuestiones estratégicas sobre si se produce un cambio sostenido en inserción social, en la iniciativa económica local o en los hábitos y actitudes de las personas.

3. EL ESTUDIO DE CASO

1. El problema

El proyecto objeto de estudio —en el marco de la iniciativa europea Equal— se plantea intervenir en el problema que tienen personas —que han dependido económicamente de sus familiares más directos— para pasar de la inactividad laboral al trabajo formal retribuido fuera de casa. Se sabe que el proceso de transición al trabajo formal y legal fuera de casa, en particular para los jóvenes y las mujeres de mediana edad requiere tiempo, tiende a ser errático y bastante distinto en cada caso individual. Se sabe que es difícil incluso para personas con un nivel de formación elevado.

Cabe esperar que una persona que empieza a trabajar —en particular las mujeres— tenga dificultades para compaginar sus obligaciones familiares con sus nuevas obligaciones laborales[5] y que necesite del tiempo de su cónyuge, que no entienda lo que se espera de ella en el trabajo o que lo haga mal y hasta sea despedida o no vea renovado su contrato. Al igual que sucede con el proceso de maduración y emancipación de los jóvenes, para una mujer de mediana edad con hijos, dejar la casa o integrarse en el mercado laboral formal es un proceso —antes que un suceso— que plantea la necesidad de un esfuerzo sostenido y con vistas al medio y largo plazo para que tenga alguna posibilidad de éxito.

Sin embargo, tanto los servicios públicos de formación ocupacional y empleo como los servicios sociales tienen dificultades para «acompañar» a sus usuarios en este proceso y adaptar los servicios «a la medida» de las personas individualmente consideradas. Los servicios de empleo están acostumbrados a trabajar con «desempleados» pero les resulta extraña la idea de las personas «inactivas» e incluso, en muchos casos, tienen dudas sobre la oportunidad de dedicar recursos públicos a la activación de personas que están en sus casas y que, de hecho, no están llamando a las puertas de la administración. Saben como derivar a sus usuarios a servicios estandarizados de formación pero tienen dificultades para adaptar sus tareas profesionales a un modelo de intervención más flexible, variado y que se desarrolle en un plazo más largo convirtiéndose, de hecho, en acompañantes o tutores de los usuarios.

Más allá del análisis del itinerario profesional y de las potencialidades de empleo, se suele ver la formación o la preparación para la búsqueda de trabajo como el fin del proceso, a pesar de que en muchas ocasiones es sólo el principio. En los servicios sociales suelen trabajar personas con vocación de ayuda a los des-

[5] La literatura sobre conciliación familia trabajo es aquí relevante. Ver Comisión Europea, 1998; Consejo de la Unión Europea, 2000; Comisión de las Comunidades Europeas, 2004, Instituto de la Mujer, 2001, Consejo Económico y Social, 2002, Escobedo y Flaquer, 2003.

favorecidos y al menos, en teoría, existen fórmulas de atención más personalizadas, si bien la finalidad de la intervención no suele ser la autosuficiencia sino la obtención de ayudas sociales que puedan suplir la falta de trabajo e ingresos.

2. Las inactivas usuarias de los servicios públicos locales de empleo

La investigación consiste en el análisis de la inactividad de mujeres de mediana edad a partir de la experiencia y las percepciones de treinta mujeres inactivas que el equipo investigador pudo seleccionar entre las participantes en un proyecto financiado por la Comisión Europea, organizado por la Diputación de Barcelona y desarrollado en unos cincuenta municipios de la provincia de Barcelona.[6] No se trata de analizar aquí los programas concretos en los que participaron estas treinta mujeres sino simplemente de aprovechar la oportunidad de poder realizar entrevistas en profundidad con ellas para intentar comprender dos cuestiones fundamentales para la política: en primer lugar, en qué consiste la inactividad y cuales son las principales categorías de inactivas —ya que por ejemplo debemos considerar que una parte de las inactivas son activas en la economía sumergida— En segundo lugar, qué peldaños componen la escalera que conecta la inactividad con el empleo formal y la medida en que familiares y servicios públicos acompañan a las mujeres inactivas y facilitan la ascensión de esta escalera.

Se aplica una metodología cualitativa que parte de la idea de dar todo el protagonismo al sujeto y a su experiencia tanto de la inactividad como de la relación con los familiares y los servicios públicos preocupados por su desarrollo profesional. La principal idea es pues la de «dar voz» a las inactivas y la de analizar su trayectoria individual desde una perspectiva temporal y dinámica. Si bien no se trata de historias de la vida en el sentido clásico ni tampoco de una reconstrucción biográfica completa (Bourdieu, 1999; Lemkow, Tejero y Torrabadella, 2000), la reconstrucción del momento biográfico en que se produce la activación permite analizar la intersección entre individuo y servicio público, entre la vida de estas mujeres y las políticas que se justifican en la intervención sobre sus problemas.[7]

Al final del texto se reproduce una tabla donde, además de incluir información sociológica básica sobre cada una de las treinta inactivas, se sintetiza información sobre experiencia laboral previa, años de inactividad, motivo de la inactividad, situación actual y perspectivas de futuro en el momento de realizarse la entrevista. Se han sustituido los nombres por números para preservar la intimidad de las personas entrevistadas.

---

[6] Proyecto «Elionor» de la Iniciativa Comunitaria Equal, cofinanciado por el Fondo Social Europeo y una Agrupación de Desarrollo representada por la Diputación de Barcelona para el período 2002-2004.

[7] Esta metodología tiene debilidades que se han podido constatar en la medida que, en primer lugar, no es posible recoger mucha información biográfica en unos pocos encuentros; en segundo lugar, al limitar la fuente de información a las mismas entrevistadas, siempre cabe la posibilidad que su relato atribuya coherencia a actos que quizás no la tuvieran en la medida en que era presentada; y, en tercer lugar, una entrevista entre un equipo de investigadores universitarios y una mujer inactiva nunca es totalmente simétrica y natural a pesar de los esfuerzos realizados para reducir esta percepción.

3. **Las lecciones aprendidas**

A. Sobre la inactividad

Las estadísticas oficiales sólo distinguen entre «activos» e «inactivos» sin entrar en el análisis de los distintos perfiles que pueden tener las personas que no tienen un «empleo formal» y que tampoco pueden ser consideradas «desempleadas» de acuerdo con la legislación. Sin embargo, cuando algunos servicios públicos, en el marco de las políticas europeas de empleo se han interesado por tratar a grupos de inactivos reales y concretos, se han encontrado con realidades muy distintas.

Las entrevistas realizadas a treinta mujeres inactivas ofrecen información que permite analizar las distintas situaciones de inactividad y las trayectorias vitales que les llevaron a esta situación. El análisis utiliza una tipología que distingue entre cuatro categorías: falsas inactivas, inactivas temporales, inactivas permanentes o resignadas e inactivas voluntarias.[8]

a. Las falsas inactivas

Existe un grupo de mujeres que nunca han estado alejadas del mercado de trabajo, pero que debido a estar empleadas en el mercado de trabajo de la economía sumergida son incluidas en la categoría de las inactivas. Las «falsas inactivas» son mujeres activas en el sentido material del término y cuentan con unos mecanismos propios de búsqueda de trabajo que las alejan por completo del mercado laboral legalmente regulado.

En primer lugar, el mercado de trabajo informal presenta ventajas para las «falsas activas» en la medida que les ofrece una flexibilidad de horarios que difícilmente pueden ofrecer los sectores más formalizados y que resulta muy atractiva para mujeres que deben compaginar su vida laboral con la familiar. Además de las ventajas que conlleva la flexibilidad horaria de determinados sectores incluidos en la economía sumergida, se observa la presencia de una trampa salarial en comparación con los sectores formales ya que el salario neto recibido en la economía sumergida es bastante superior al percibido en el sector formal al verse libre de cotizaciones sociales e impuestos.

Los sectores en los que predomina la actividad de las falsas inactivas son inequívocamente aquellos relacionados con las tareas domésticas, el textil y la hostelería, aunque éstos últimos se pueden considerar temporales y estacionarios y dependientes de flujos de producción o de la temporada alta en el caso de la hostelería.

En el caso de las «falsas inactivas», las mismas causas que favorecen su incorporación a los sectores informales son las que las mantienen dentro de ellos: la demanda de trabajo y la flexibilidad inherente en los trabajos de asistencia doméstica junto con su elevada remuneración en comparación con los sectores formales de la economía mantienen a estas mujeres en el sector informal. En el mismo sentido, la flexibilidad horaria facilita la conciliación de la vida familiar y laboral, lo que es valorado por estas mujeres ante

[8] Esta no es una tipología oficial pero si ha sido utilizada por los gestores locales de los programas inducidos por redes de cooperación como la desarrollada por la Diputación de Barcelona. Ver Domínguez, 2004.

el coste de las guarderías privadas, de los comedores escolares y de las canguros o la necesidad de atención periódica de un familiar con problemas de salud, ya sean niños o abuelos. Cuando se producen este tipo de situaciones, son las mujeres las que se encargan de realizar estas tareas por disponer de horarios adaptados lo que, a su vez, les impide ocuparse en la economía formal.

El acceso a la economía formal es percibido como complicado y difícil por entender que la edad y las carencias formativas o de especialización lo hacen casi imposible. La consecuencia de esta percepción sobre sus debilidades formativas es su deseo de aprender y de formarse a pesar de que tampoco estén muy seguras de que vayan a poder hacerlo por la percepción que tienen de su situación personal y familiar. La obtención de un mayor grado de formación profesional es percibido como una condición necesaria para sentirse seguras e intentar el acceso a puestos de trabajo en condiciones dignas dentro del sector formal.

b. Las inactivas temporales

Esta categoría ilustra el paso, para un número importante de mujeres, desde una situación normalizada de integración en el mercado laboral formal, normalmente desde edades muy tempranas, a una situación de desempleo primero y de inactividad después, que se justifica por razones familiares y en particular vinculadas a la maternidad.

El distanciamiento del mercado de trabajo es en principio voluntario y temporal si bien las entrevistas revelan un número elevado de casos percibidos como de «*mobbing*» con consecuencias psicológicas, así como una contradicción entre el deseo de volver al trabajo una vez transcurrido el tiempo en que los motivos del alejamiento inicial pierden relevancia, y la constatación de las crecientes dificultades objetivas para conseguir un puesto de trabajo y desarrollar una jornada completa dadas las obligaciones familiares.

Resulta evidente que la maternidad condiciona la vida laboral de muchas mujeres. Ello es aún más evidente en aquellos casos de mujeres trabajando en situaciones de mayor inestabilidad laboral, que no cuentan con el apoyo de familiares y que no tienen recursos para pagar guarderías privadas, comedores escolares y canguros. Por otra parte, el deseo de permanecer junto a los hijos los primeros años de vida aporta sentido y coherencia a la decisión de mantenerse en una inactividad que se estima temporal y por un período de tiempo razonable.

Las entrevistas recogen asimismo varios casos de complicaciones durante el embarazo, de problemas de salud propios y ajenos, de cuidado periódico de uno de los padres en situación de viudedad, que suponen otros tantos motivos coherentes para romper el vínculo laboral e iniciar una situación de desempleo o inactividad. En los casos de «mobbing,» son importantes las secuelas en la medida que las afectadas se sienten aún más alejadas del mercado de trabajo y perciben la necesidad de rearme y recuperación psicológica ante una eventual vuelta al mundo exterior de las empresas y del trabajo fuera de casa.

Las entrevistas permiten constatar actitudes y comportamientos heterogéneos ante situaciones objetivamente similares.

Sin embargo, se puede afirmar que existe una voluntad generalizada entre las inactivas temporales de recuperar el trabajo fuera de casa y de participar en el sostenimiento económico de la familia ante la insuficiencia o las limitaciones al bienestar de las familias que supone la percepción de un único salario.

La percepción de un período de tiempo prudencial de inactividad tras la primera maternidad se sitúa entre 2 y 4 años. Pasado este tiempo, algunas entrevistadas —pocos casos relativos— se consideran a sí mismas «desempleadas» y buscadoras activas de empleo con algunas diferencias en función del entorno en el que habitan y de sus circunstancias personales y familiares.

En los casos de una segunda maternidad, el período de inactividad se prolonga notablemente. Uno de los factores más determinantes del no retorno prolongado al mercado laboral es la presencia de un segundo hijo. Así, se puede afirmar que si la primera maternidad es un factor determinante de la inactividad temporal, la segunda maternidad —entre las 30 entrevistadas, sólo hay cuatro mujeres con más de dos hijos, dos de ellas son inmigrantes— es determinante de la inactividad prolongada y en algunos casos permanente.

c. Las inactivas permanentes o resignadas

Esta categoría coincide prácticamente con el grupo de mayor edad y de mayor tiempo de inactividad. La media para todas las usuarias del programa de la Diputación de Barcelona —más de 2.300 usuarias— se situaba en torno a los 11 años de inactividad. Tras más de una década de inactividad y con edades que se consideran avanzadas por parte de las mismas entrevistadas, su propia percepción de las posibilidades de éxito en el mercado de trabajo es mucho menor y el recurso a la economía informal o a resignarse a su condición de inactivas es más probable.

Sin embargo, estas mujeres quieren trabajar y ven el trabajo como un deber social. Tal como se ha podido comprobar en las entrevistas, en el momento en que las necesidades de la familia empiezan a decrecer (hijos en secundaria, fallecimiento de ancianos a cargo…), las inactivas involuntarias inician un proceso de búsqueda de salidas a su situación. Prueba de ello es que la mayoría de las entrevistadas se acercan a los servicios municipales de promoción económica y no necesitan se reclutadas por los técnicos que llevan los programas europeos o locales de empleo. Sin embargo, este regreso al mercado de trabajo conlleva numerosas dificultades, tanto de orden objetivo (edad, falta de formación, comportamiento de las empresas, coste de guarderías, comedores escolares y canguros) como de orden subjetivo (baja autoestima, miedo a las consecuencias de la menor dedicación a la familia, actitudes de los familiares directos).

d. Las inactivas voluntarias

Las entrevistas también ponen de relieve que algunas mujeres optan de hecho por la inactividad. Ante circunstancias similares, algunas mujeres expresan claramente y de forma consciente en las entrevistas su preferencia por trabajar «dos o tres horas por las mañanas» o simplemente por no trabajar. Entre las inactivas

voluntarias hay algunas mujeres que no accedieron nunca al mercado de trabajo formal. La mayoría sin embargo, habiendo entrado de manera temporal dentro de la inactividad (normalmente por maternidad), transforman su estado temporal en permanente, abandonando de hecho el mercado de trabajo.

Las razones que explican esta situación deben buscarse una vez más en la cultura que continúa perpetuando los roles de género y en las condiciones de trabajo que ofrece el mercado a estas mujeres. La visión tradicional de la familia conlleva tanto la asunción por el hombre del papel protagonista en el ámbito público y laboral como la atribución de la responsabilidad a la mujer por todo lo que tenga que ver con el ámbito reproductivo y asistencial dentro de la familia. El otro factor explicativo de la ubicación en esta categoría de gran número de inactivas es el bajo nivel de los salarios a los que podrían aspirar, la deficiente organización del trabajo desde la perspectiva de poder llevar una vida familiar que pueda calificarse como tal y el hecho de no necesitar los ingresos al tener las necesidades mas básicas cubiertas por la actividad de los esposos o por prestaciones que no son de jubilación.

En los casos de necesidad económica grave o sobrevenida —por una situación de desempleo del cónyuge o por la ruptura del matrimonio o unión extramatrimonial— el primer recurso de estas mujeres es la realización de trabajos en la economía sumergida combinada con la percepción de ayudas sociales del tipo de renta mínima.

B. Sobre la intervención pública ante la inactividad

La intervención pública está claramente justificada cuando se produce la renuncia al trabajo por resignación. Esta categoría debería centrar el mayor esfuerzo de los servicios públicos en la medida que la lógica de las políticas activas de empleo es precisamente la de que las personas que quieren trabajar puedan desarrollar su derecho individual al trabajo. Desde la perspectiva política, existe un amplio consenso en que los servicios públicos deberían contribuir a remover los obstáculos objetivos y externos que, con el paso del tiempo, ponen cada vez más difícil la integración laboral de algunas personas.

Los casos de falsa inactividad e inactividad temporal también deberían ser objeto de la atención de los servicios públicos si bien los primeros responden a una lógica de legalización y mejora de las condiciones de trabajo, mientras que los segundos, en los que la inactividad responde a una opción en principio personal, presentan otro tipo de riesgo: que con el paso del tiempo se conviertan en casos de inactividad permanente y por tanto de renuncia por desistimiento. Respecto a este último grupo, sin embargo, una parte de la sociedad podría discutir la utilización de recursos públicos para activar a personas que aparentemente han optado por la inactividad temporal en beneficio del cuidado de sus familias.

C. Sobre la eficacia de los servicios públicos

La otra cuestión planteada es si los servicios públicos, y en particular los de proximidad, son capaces de llevar a

cabo y con eficacia la misión de acompañar a mujeres de mediana edad desde la inactividad formal a la actividad laboral fuera de casa.

Los programas europeos de empleo permiten a las administraciones locales contratar a técnicos de ocupación y crear por tanto estructuras *ad hoc* para la atención de sus beneficiarios. Ello supone una ventaja relativa respecto a otros programas que utilizan recursos de su estructura administrativa en la medida que aquellos pueden ser seleccionados en función de su motivación e interés por el problema de la inactividad femenina.

Pero a la luz de las entrevistas con las treinta usuarias, no todas las iniciativas locales tienen una filosofía definida respecto al modelo que deben seguir ni tampoco una teoría de acción sobre como actuar. A ello no contribuye la ambigüedad en los objetivos. Desde la perspectiva de los poderes públicos, los objetivos se suelen interpretar como «conseguir un trabajo para la participante.» Pero a menudo se acaba adoptando como objetivo la idea de «activar laboralmente» a las beneficiarias de los servicios de proximidad. El trabajo es el fin, si bien, dadas las dificultades para encontrar empleos, a menudo los servicios públicos acaban *justo antes* de que las beneficiarias empiecen su trabajo y de hecho en la mayoría de los casos no saben qué les ha pasado a las beneficiarias de sus programas después de la finalización de los talleres.

La mayor innovación en los servicios de empleo ha sido la adopción de una filosofía de gestión que busca un mayor grado de proximidad (desde la administración local), personalización (diagnósticos caso por caso, terapia de grupo) y flexibilidad (derivación a los cursos ocupacionales). Sin embargo, en la práctica, los servicios que se ofrecen a las usuarias no ofrecen grandes variaciones respecto a un modelo básico que consiste en una combinación de los siguientes elementos: entrevista individualizada, activación psicológica en grupo, introducción a las tecnologías de la comunicación y de la información, preparación de *currículums* y entrevistas de selección junto con la posibilidad de asistir a algún curso ocupacional y, en algunos casos, realizar prácticas en empresas —una opción muy interesante para las beneficiarias ya que les permite familiarizarse con un entorno de trabajo real y, en algunas ocasiones, obtener un empleo—.

Las entrevistadas, en términos generales satisfechas con los servicios recibidos, apuntan sin embargo a la principal debilidad de este modelo al señalar que los programas públicos terminan justo en el momento que más necesitan su apoyo, cuando empiezan a buscar trabajo y, en algunos casos, a trabajar. La principal dificultad y el principal temor para estas mujeres consiste en encontrar un puesto de trabajo y una vez conseguido, en mantenerse trabajando, debido a factores como su falta de experiencia, el comportamiento oportunista de las empresas, los horarios de tarde y noche, el coste de los servicios —en particular comedores escolares y canguros— y las presiones familiares ante los cambios en el tiempo que pueden dedicar a la familia.

Lógicamente, cabe esperar que las mujeres inactivas se sientan satisfechas por el hecho de ser atendidas por técnicos especialmente contratados para tener cuidado de su activación laboral.

Sumario

Frente al olvido o la soledad en casa, las usuarias de este servicio público gratuito se sienten felices de entrar en contacto con otras mujeres con problemas similares y de iniciar una terapia de grupo motivadora que las anime a desarrollar un nuevo estilo de vida más independiente.

Pero el éxito en su proceso de inserción se vería incrementado si los servicios públicos cambiaran de orientación y apoyaran de manera sostenida a las demandantes de empleo a partir del momento en que se lanzan a la calle a buscar trabajo. En este momento crítico de establecer contacto con potenciales empleadores, de presentarse a pruebas de selección y de empezar a trabajar, pocos servicios públicos acompañan a sus usuarias en la medida que entienden que ya han completado el ciclo previsto. Todo ello a pesar de que éste es el momento crítico en el que se produce el abandono de personas que no cuentan con elevados niveles de apoyo material (por horarios incompatibles, cuidado de familiares directos, comidas al mediodía de hijos menores) y psicológico (aguantar las pruebas de selección hasta el final, mantener un elevado nivel de dedicación al trabajo con el objeto de convencer al empresario, sentir que la familia no va notar el cambio o que vale la pena sacrificar algunos pequeños detalles a cambio de un trabajo regular). El acompañamiento continuado supondría reforzar la filosofía de progresión escalonada hacia la independencia económica y apoyo a largo plazo frente a la inercia de la administración pública tradicional que se caracteriza por su visión estandarizada y desde la distancia en la definición y el tratamiento de las personas.

Las entrevistas también pusieron de relieve como una de las cuestiones más difíciles consiste en encontrar la justa medida entre la voluntad de trabajar y ganar dinero de forma inmediata y la oportunidad de cumplir con un programa de formación, prácticas y empleo a tiempo parcial que permita una progresión profesional. En estos procesos de negociación, es importante que los servicios públicos cuenten con recursos adicionales para cubrir comedores escolares o canguros y que las mujeres sean capaces de crear un entorno familiar que les permita ganar tiempo para sí mismas, cumplir sus compromisos, y que tanto ellas como sus familiares se adapten a la nueva situación.

4. CONCLUSIONES SOBRE LA EVALUACIÓN DE POLÍTICAS DE EMPLEO EUROPEAS

La evaluación de un proyecto de empleo de mujeres inactivas como el utilizado en este artículo podría haberse limitado a reproducir y explicar los resultados principales en términos de beneficiarios directos e indirectos, actividades organizadas, acciones formativas, contactos realizados con personas inactivas, entrevistas y tutorías, actos de difusión, seminarios y jornadas, folletos y carteles informativos o incluso aparición en medios de comunicación audio visuales y de prensa escrita. La justificación del proyecto se lograría con una gestión ordenada, conforme a las previsiones temporales, y con unos índices de actividad y de beneficiarios razonables en función del nivel de financiación pública recibida.

De acuerdo con esta lógica, la evaluación cumple la función de aportar evidencias sobre la corrección de los procesos de implementación locales y del número

de actividades y beneficiaros implicados en los proyectos, lo que permite a la administración europea distribuir los recursos en base a criterios de racionalidad política y administrativa y continuar con el ciclo de gestión de proyectos. En este contexto, no es previsible que el cambio de política —*policy shift*— se produzca como consecuencia del análisis o la evaluación ni que ésta tenga influencia alguna en el proceso de decisión política.

Sin embargo, con una metodología sencilla como las entrevistas realizadas a treinta personas inactivas participantes en programas locales de empleo, se ha podido señalar algunas cuestiones —sobre las categorías de inactivas o sobre la necesidad de apoyo continuado a partir del momento en que la administración entiende terminados los servicios— que habrían pasado desapercibidas en una evaluación estándar. Se podría afirmar pues que de la misma manera que las usuarias de estos servicios precisan servicios con mayores niveles de personalización, flexibilidad y continuidad a largo plazo para avanzar en la progresión que lleva a su plena integración laboral, la evaluación también debe superar el modelo burocrático de control de las actividades y adaptarse a cada programa para, a partir de la definición de sus resultados principales y de la estrategia que persigue para alcanzarlos, desvelar realmente la medida en que cumple con sus promesas.

Ciertamente, no cabe esperar que la evaluación de un proyecto tenga una influencia directa en la formulación de nuevas políticas pero la coincidencia de varios análisis en señalar las principales debilidades de una forma de proceder sí puede acabar haciendo mella en los responsables políticos y administrativos. Una vez alcanzada una cierta capacidad de control en la ejecución presupuestaria, en la contabilidad analítica o en la gestión de actividades, los responsables políticos necesitan información cualitativamente rica que les permita introducir reformas. Sin embargo, la evaluación suele producir información bastante anodina en lugar de cuestionar los objetivos y la lógica de las intervenciones para alimentar debates de mayor relevancia teórica y política.

Sumario

Cuadro resumen mujeres inactivas. Entrevistas realizadas por Anna Antó, Vanesa de la Fuente y Manuel Pérez.

Nombre*	Edad	Estado civil	Hijos	Educación	Experiencia	Años inactiva	Motivo de la inactividad	Situación actual	Perspectiva de futuro	Fecha entrevista	Tipología*
1.	45	Casada	2 de 22 y 19 años	Bachillerato	Dependienta durante 8 años	20	Cierre de empresa / familia	Estudios de recepcionista	Trabajar	27/10/2004	Inactivas permanentes o resignadas
2.	45	Casada	2 de 12 y 10.	1º Psicología	Administrativa durante 16 años	12	Mobbing / familia.	Trabaja (contrato de inserción laboral INEM)	Trabajar	6/10/ 2004	Inactivas permanentes o resignadas
3.	35	Casada	1 de 5 años	F.P.II administrativa	Amplia experiencia como administrativa	2	Mobbing	Estudios de comunicación.	Trabajar	27/10/2004	Inactiva Temporal
4.	45	Divorciada y viviendo en pareja	2 de 9 y 8 años.	E.G.B	Dependienta, hostelería, conductora, verificadora…	12	Familia	Estudios de informática.	Formarse	2/11/ 2004	Inactiva Voluntaria
5.	38	Casada	2 de 16 y 11 años.	F.P.I Peluquería	Peluquera, logística y monitora escolar	5	Familia	Estudios de informática.	Formarse	2/11/ 2004	Inactiva Voluntaria
6.	33	Casada	1 de 4 años.	F.P Jardín de Infancia	Monitora de guardería, cajera de supermercado	4	Familia	Estudios de administrativa	Trabajar	13/10/2004	Inactiva temporal
7.	45	Casada	1 de 11 años	Graduado Escolar	Ayudante consulta médica, cosméticos, dependienta	2	Despido / Familia	Está igual que antes	Trabajar	2/11/2004	Inactiva temporal

Cuadro resumen mujeres inactivas. Entrevistas realizadas por Anna Antó, Vanesa de la Fuente y Manuel Pérez.

Nombre*	Edad	Estado civil	Hijos	Educación	Experiencia	Años inactiva	Motivo de la inactividad	Situación actual	Perspectiva de futuro	Fecha entre-vista	Tipología*
8.	47	Casada	2 de 20 y 15	Bachillerato elemental	Auxiliar de farmacia durante 11 años	20	Familia	Trabaja	Trabajar	28/10/ 2004	Inactivas permanentes o resignadas
9.	40	Separada	2 de 12 y 9 años	Graduado Escolar	Secretaria, asistenta, economia sumergida	9	Familia	Está igual que antes	Trabajar	28/10/ 2004	Inactivas permanentes o resignadas
10.	48	Casada	2 de 24 y 21 años	Graduado Escolar	Bordadora de Máquinas Electrónicas	20	Familia	Trabaja / monitora de comedor escolar	Formarse y trabajar	4/10/ 2004	Inactivas permanentes o resignadas
11.	36	Casada	2 de 13 y 8 años	Graduado Escolar	Operaria industrial, planchadora, almacén, dependienta...	0 (Economia sumergida)	Problemas médicos del hijo	Trabaja en economía sumergida y estudios	Formarse y trabajar	4/10/ 2004	Falsa Inactiva
12.	42	Casada	2 de 11 y 3 años	Graduado Escolar	Administrativa y operaria textil	15	Salud y Familia	Está igual que antes	Formarse y trabajar a tiempo parcial	4/10/ 2004	Inactiva voluntaria
13.	34	Casada	2 de 7 y 3,5 años	Auxiliar Admtva	Informática y administración	3,5	Familia	Está igual que antes	Trabajar si lo necesita	14/10/ 2004	Inactiva voluntaria
14.	37	Casada	2 de 5 y 3 años	F.P.I	Administrativa, dependienta y operaria	5 (economía sumergida)	Salud y Familia	Está igual que antes	Formarse y trabajar a tiempo parcial	14/10/ 2004	Falsa Inactiva

Sumario

Cuadro resumen mujeres inactivas. Entrevistas realizadas por Anna Antó, Vanesa de la Fuente y Manuel Pérez.

Nombre*	Edad	Estado civil	Hijos	Educación	Experiencia	Años inactiva	Motivo de la inactividad	Situación actual	Perspectiva de futuro	Fecha entre-vista	Tipología*
15.	40	Casada	3 de 15, 11 y 9 años	Graduado Escolar	Operaria en fábricas de boradados	15	Familia.	Trabaja a tiempo parcial / geriatríco	Trabajar a tiempo parcial	20/10/ 2004	Inactivas permanentes o resignadas
16.	38	Casada	3 de 12, 9 y 6 años.	1º Curso de Psicología en Nicaragua.	Trabaja en negocio familiar en Nicaragua. En España, cuida niños	4 (economia sumergida)	Inmigrante	Trabaja / cuidadora de niños	Volver a su país para dirigir el negocio familiar	6/10/ 2004	Falsa Inactiva
17.	31	Casada	1 de 4 años	E.G.B	Dependienta, canguro, operaria, cartera…	0	No encuentra trabajo estable.	Trabaja	Trabajar	18/10/ 2004	Falsa Inactiva
18.	51	Casada	2 de 14 y 11 años	Licenciada en Psicología. Curso de Dirección de Residencias	Comunicadora terapéutica	1,5	Familia	Está igual que antes	Trabajar a tiempo parcial	20/10/ 2004	Inactiva temporal
19.	41	Casada	2 de 8 y 5 años.	Graduado Escolar	Operaria industrial	11	Familia	Está igual que antes	Formarse.Trabajar si le gusta el trabajo	26/10/ 2004	Inactivas permanentes o resignadas
20.	34	Viuda	0	Estudios en Empresaria-les y Márketing	Ventas, marketing y comercio exterior	1	Calidad de vida	Está intentando montar una empresa	Trabajar	26/10/ 2004	Inactiva temporal
21.	40	Casada	2 de 11 y 8 años.	B.U.P inacabado	Dependienta, operaria textil, monitora	8	Rescisión del contrato / familia	Trabaja a tiempo parcial	Trabajar	26/10/2004	Inactivas permanentes o resignadas

Cuadro resumen mujeres inactivas. Entrevistas realizadas por Anna Antó, Vanesa de la Fuente y Manuel Pérez.

Nombre*	Edad	Estado civil	Hijos	Educación	Experiencia	Años inactiva	Motivo de la inactividad	Situación actual	Perspectiva de futuro	Fecha entre-vista	Tipología*
22.	35	Separada	4 de 16, 14, 11 y 5 años	Asiste a la escuela de adultos	Operaria sector textil	2	Inmigrante	Estudios de Formación Ocupacional	Trabajar	12/1/2004	Inactiva Temporal
23.	36	Soltera	1 de 2 años	FP I Auxiliar Clinico	Geriátricos y disminuidos psícicos	2	Salud y Familia	Estudios de Formación Ocupacional	Formarse y trabajar	12/1/2004	Inactiva temporal
24.	38	Casada	2 de 7 y 4años	EGB	Operaria textil	7	Familia	Formación	Trabajar	12/1/2004	Inactivas permanentes o resignadas
25.	43	Casada	1 de 18 años	COU y auxiliar guardería	Encargada cafeteria familiar	3	Cambio de lugar de residencia	Está igual que antes	Formarse pero no trabajar	26/10/2004	Inactiva voluntaria
26.	52	Casada	1 de 22 años	EGB	Operaria fábrica de punto, dependienta, auxiliar	3	Familia	Está igual que antes	Trabajar si le gusta el trabajo	12/1/2004	Inactiva voluntaria
27.	55	Casada	3 de 33, 32 y 30 años.	Bachillerato Superior y Técnica Publicitaria.	Aux. admtva, comercial y cocinera	10	Salud y Familia	Formación	Trabajar si le gusta el trabajo	4/11/2004	Inactiva voluntaria
28.	43	Viuda	1 de 16 años	Licenciada en Contabilidad y Finanzas	Gerente empresa comercio exterior	7	Familia	Formación	Trabajar si le gusta el trabajo	7/10/2004	Inactiva voluntaria

Sumario

Cuadro resumen mujeres inactivas. Entrevistas realizadas por Anna Antó, Vanesa de la Fuente y Manuel Pérez.

Nombre*	Edad	Estado civil	Hijos	Educación	Experiencia	Años inactiva	Motivo de la inactividad	Situación actual	Perspectiva de futuro	Fecha entre-vista	Tipología*
29.	55	Casada	2 de 30 y 24 años	Graduado escolar	Modista	0 (economía sumergida)	Familia	Trabaja	Trabajar	7/10/2004	Falsa Inactiva
30.	45	Casada	2 de 16 y 11 años	Graduado escolar	Aprendiz de peluquería y dependienta	16	Familia	Está igual que antes	Nada	7/10/2004	Inactiva Voluntaria

Fuente: Elaboración propia.

Sumario

REFERENCIAS BIBLIOGRÁFICAS

BOURDIEU, PIERRE (DIR.) (1999).:*La miseria del mundo*, México: Fondo de Cultura Económica de Argentina.

COMISIÓN EUROPEA CE-V-2-98-008-C (1998): *Conciliación del trabajo y la vida familiar en Europa.* Comisión Europea, DGV.

COMISIÓN DE LAS COMUNIDADES EUROPEAS COM 115 FINAL (2004): *Informe anual sobre igualdad de oportunidades entre hombres y mujeres en la Unión Europea en 2004.* Bruselas: Comisión de las Comunidades Europeas.

CONSEJO ECONÓMICO Y SOCIAL (2002): *Segundo Informe sobre la situación de las mujeres en la realidad sociolaboral española.* Madrid: Consejo Económico y Social.

CONSEJO DE LA UNIÓN EUROPEA (2000): *Articulación de la vie familiale et la vie professionnelle* SOC 373/12577, Bruselas, Consejo de la Unión Europea.

DENTE, BRUNO (2003): «La política de la evaluación de políticas» en *La evaluación de las políticas regionales,* Fundació Carles Pi i Sunyer, d'Estudis Autonòmics i Locals, Barcelona.

DOMÍNGUEZ, CARMEN (2004): *Conciliación y empleabilidad*, Barcelona: Diputación de Barcelona.

ESCOBEDO, ANNA Y LLUÍS FLAQUER (2003): «Cambios familiares, políticas familiares y de conciliación de la vida laboral y personal» en *Tercer Informe Randstad. Calidad del trabajo en la Europa de los Quince: Las políticas de conciliación.* Barcelona, Instituto de Estudios Laborales de ESADE.

GULJT, I Y J. GAVENTA (1998): *Participatory Monitoring and Evaluation*, Institute of Development Studies, University of Sussex, Brighton.

HATRY, H (1999): *Performance measurement: getting results*, The Urban Institute, Washington DC.

INSTITUTO DE LA MUJER (2001): *Guía de buenas prácticas para conciliar la vida familiar y profesional,* Madrid: Instituto de la Mujer, Ministerio de Trabajo y Asuntos Sociales.

KELLEHER, JOHN; SARAH BATTERBURY, ELLIOT STERN (1999): «*The Thematic Evaluation of the Partnership Principle*, The Tavistock Institute.

LEMKOW, LUÍS, ELISABETH TEJERO Y LAURA TORRABADELLA (2000): *Costos biogràfics en la lluita pel benestar*, Barcelona: Editorial Mediterrània.

MAIRATE, ANDREA, (2003): «La evaluación de los fondos regionales: problemas y retos de futuro,» en *La evaluación de las políticas regionales,* Fundació Carles Pi i Sunyer, d'Estudis Autonòmics i Locals, Barcelona.

WEISS, CAROL H. (1998, 2ª edición): *Evaluation* Prentice Hall Upper Saddle River, New Jersey.

WEISS, CAROL H. (2000): «Theory-based evaluation: theories of change for poverty reduction» en O. Feinstein y R. Picciotto (eds.) *Evaluation and poverty reduction*, The World Bank, Washington DC.

WHOLEY, J.S. (1983): *Evaluation and effective public management*, Boston, Little Brown.

Sumario

El margen de maniobra presupuestario de las Comunidades Autónomas. Método de análisis y aplicación a la Comunidad Autónoma de Aragón

Todos los gobiernos aspiran a obtener la máxima discrecionalidad o margen de maniobra en la gestión de su presupuesto. Sin embargo, esa discrecionalidad en la gestión presupuestaria se ve limitada inevitablemente, tanto externamente —decisiones de otros niveles de gobierno de ámbito superior—, como internamente —decisiones propias pretéritas que crean obligaciones de gasto—. El trabajo que presentamos nace del interés por conocer cuál es y cómo ha evolucionado la discrecionalidad o margen de maniobra presupuestario de las comunidades autónomas. Para ello, se propone un modelo de análisis basado en la clasificación de las operaciones de ingreso y gasto y en la explotación de información estrictamente presupuestaria. Se toma como caso de estudio la actividad presupuestaria en la Comunidad Autónoma de Aragón entre 1986 y 2000. Los resultados ponen de manifiesto la notable potencialidad analítica de esta línea de investigación.

Gobernu orok bere aurrekontuan eragiketak egiteko ahalmen handiena lortu nahi du. Hala eta guztiz ere, aurrekontuaren kudeaketan duen eskumen hori mugatuta dago nahitaez, bai kanpotik —goragoko aginte-mailen erabakien ondorioz—, bai barrutik —gastu-obligazioak dakartzaten iraganeko erabakien ondorioz—. Autonomia-erkidegoen eskumena edo aurrekontuetan eragiketak egiteko aukera zein den eta nola garatu den jakiteko interesak eraman gaitu lan hau aurkeztera. Horretarako, diru-sarreren eta gastuen eragiketak sailkatu eta aurrekontuari buruzko informazioa erabiltzean oinarritutako azterketa-eredua proposatu dugu. Aragoiko autonomia-erkidegoak 1986. eta 2000. urteen artean izandako aurrekontu-jarduera hartu da aztergaitzat. Emaitzek ikerketa-lerro honek analisiak egiteko balio duela agerian utzi dute.

Each government aspires to obtain the maximum capacity of manoeuvring with its budget. But, that discretionary act in the budgetary management is inevitably limited, both externally- decisions of other levels of superior scope government—, as well as internally – former proper decisions that create obligations of expense. The work that we present came out from the interest to know what is and how had developed the discretionary act or the budgetary manoeuvring margin of the Autonomous Communities.
An analysis model is proposed which is based on the classification of the income and expense operations and on the management of strictly budgetary information. The case to be studied is the budgetary activity in the Autonomous Community of Aragón between 1986 and 2000. The results point out to the important analysing potentiality of this line of investigation.

Ramón Barberán Ortí
Universidad de Zaragoza

ÍNDICE

1. Introducción
2. Modelo de análisis
3. Resultados de la aplicación a la Comunidad Autónoma de Aragón
4. Conclusiones
Referencias bibliográficas

Palabras clave: presupuesto, disciplina presupuestaria, margen de maniobra presupuestario, comunidades autónomas, Aragón.

N.º de clasificación JEL: H72

1. **INTRODUCCIÓN**

La elaboración y aprobación del presupuesto es una de las principales tareas anuales a las que se enfrenta cualquier administración pública, ya que implica establecer las pautas a las que se deberán ajustar sus intervenciones financieras en el futuro inmediato, desde los grandes proyectos de inversión a las pequeñas partidas de gasto corriente, pasando por toda la amplia diversidad de transferencias. Ninguna administración, por pequeño que sea su tamaño, puede sustraerse a la sensación de que en el proceso presupuestario reside buena parte de su capacidad para decidir sobre la asignación de recursos y la redistribución de la renta. Sin embargo, la capacidad de cualquier administración de influir sobre el volumen global del presupuesto y sobre su estructura y destino concreto varía muy sustancialmente según sus competencias en materia de ingresos, sus responsabilidades en materia de gastos y los compromisos adquiridos en el pasado.

Este trabajo se ocupa, precisamente, de la medición del grado de discrecionalidad de que gozan los gobiernos regionales durante el proceso presupuestario (elaboración y ejecución del presupuesto) para decidir sobre el destino de los fondos de que disponen en cada ejercicio: grado de discrecionalidad al que denominamos «margen de maniobra presupuestario».

Aunque el objeto de estudio guarda relación con la autonomía financiera de las

comunidades autónomas, entendida como su capacidad para tomar decisiones sobre el volumen, la estructura y el destino de sus ingresos, no es equivalente, pues aquí asumimos como un dato el volumen y la estructura de los ingresos, y nos centramos exclusivamente en la capacidad para decidir el destino de los mismos. Por tanto, el trabajo no analiza lo que los gobiernos regionales podrían hacer si ejercitasen las posibilidades —considerablemente amplias en la actualidad— que la legislación sobre su sistema de financiación les permite (por ejemplo, la modificación de tipos en numerosos tributos cedidos, la imposición de recargos o el recurso al crédito), sino lo que pueden hacer con los ingresos de que efectivamente disponen en cada momento.

Utilizamos el término «gobierno regional» en un sentido amplio, más allá de su estricto significado jurídico, ya que en él incluimos tanto la acción del poder ejecutivo como la del legislativo. No en vano, la asignación de fondos se dilucida de modo especial en la fase de elaboración y aprobación del presupuesto (fijación del importe y distribución por programas y conceptos de los créditos para gastos), fase en la que intervienen ambos poderes. Además, el ejecutivo puede adoptar decisiones de asignación relevantes en la fase de ejecución: introduciendo modificaciones presupuestarias mediante cualesquiera de las vías disponibles al efecto; comprometiendo los créditos dentro de los conceptos de gasto aprobados (algunos de los cuales, como ciertos programas de transferencias, permiten una amplia discrecionalidad); o, simplemente, modulando el grado de ejecución de los créditos o su desarrollo temporal, especialmente en el caso de las inversiones y las transferencias.

El objetivo de este trabajo es diseñar un modelo de análisis que permita calcular el margen de maniobra presupuestario de cualquier gobierno regional en cualquier ejercicio pasado o presente, a partir de los datos que, sobre las operaciones de ingreso y gasto ejecutadas o previstas, se encuentran registrados en su cuenta de liquidación o en su presupuesto. Esto requiere un esfuerzo notable en la clasificación de las operaciones presupuestarias, de modo que sea posible distinguir la parte del gasto realizado que depende de decisiones autónomas adoptadas dentro del proceso presupuestario, de la parte que se ve afectada por condicionantes externos —decisiones de otros niveles de gobierno de ámbito superior— o por condicionantes internos —decisiones pretéritas del propio gobierno que crean obligaciones de gasto—.

Se pretende, por tanto, obtener una herramienta que permita medir qué se puede esperar realmente del proceso presupuestario en cuanto a la introducción de variaciones sustanciales entre ejercicios en la política de gasto. El conocimiento de esta cuestión tiene un considerable interés en todo momento, pero de modo muy especial cuando se producen cambios de gobierno, ya que, si el margen de maniobra realmente disponible es muy bajo, el nuevo gobierno se puede encontrar con serias dificultades para aplicar su programa. En tal caso, frente a las alternativas de renunciar a las políticas propias o incrementar la presión fiscal, ese gobierno puede optar por alternativas menos ortodoxas, como el recurso al endeudamiento (pasivos financieros) o la realización de operaciones extra-presupues-

Sumario

tarias. Así, la insuficiencia del margen de maniobra puede inducir a la indisciplina presupuestaria, lo que cobra particular relevancia en un contexto como el actual en que la estabilidad presupuestaria se sitúa en el primer plano de los objetivos de las administraciones públicas españolas y europeas. Junto a estas implicaciones de política fiscal, destacan las de gestión pública, ya que la flexibilidad en la adopción de decisiones presupuestarias constituye una condición necesaria para un funcionamiento eficiente de la administración y para la eficacia de los gobiernos.

Con arreglo a lo expuesto, el planteamiento del trabajo es básicamente metodológico, ya que no existe ninguna investigación previa que pueda adoptarse como punto de partida[1]. En su desarrollo se toma como referencia el caso de la Comunidad Autónoma de Aragón entre 1986 y 2000, al objeto de dar soporte a la clasificación de las operaciones de ingreso y gasto y, también, para ilustrar las posibilidades que ofrece en la práctica este modelo de análisis de la actividad presupuestaria de los gobiernos regionales.

Tras esta introducción, el trabajo se estructura en tres secciones. En la primera se desarrolla el modelo de análisis, es decir, se aborda la tarea de definición e instrumentación de la medida del margen de maniobra. En la segunda se presentan los resultados de la aplicación a la Comunidad Autónoma de Aragón. La tercera y última ofrece conclusiones.

2. MODELO DE ANÁLISIS

El modelo de análisis que proponemos para medir el grado de discrecionalidad o margen de maniobra presupuestario de los gobiernos regionales descansa en la clasificación de los recursos y empleos. Al objeto de reflejar del modo más preciso posible la realidad, sin perder en operatividad, distinguimos cuatro grados de discrecionalidad en las decisiones presupuestarias: 1) discrecionalidad plena; 2) discrecionalidad alta; 3) discrecionalidad baja; 4) discrecionalidad nula. Esta clasificación la aplicamos sucesivamente a los recursos y a los empleos, aunque, como veremos, el verdadero margen de maniobra presupuestario se sustancia en el grado de discrecionalidad de los empleos[2].

En el caso de los recursos atendemos a la mayor o menor capacidad del gobierno regional para fijar su destino final concreto. En el de los empleos nos fijamos, tanto en las limitaciones impuestas por sus fuentes de financiación como, adicionalmente, en las derivadas del pro-

[1] El único trabajo empírico con un objetivo similar, referido al gasto del Estado en 1989 y 1990, es el de Gómez J., García A., Jorrín J. A.; Martín J. V. y Trueba J. C. (1990). Se trata de un trabajo pionero y riguroso, pero que, sin embargo, se centra exclusivamente en el estudio de los fundamentos jurídicos de cada tipo de gasto, con lo que queda un tanto alejado de nuestra perspectiva más económica y, sobre todo, resulta metodológicamente muy poco operativo si se piensa en la extensión del estudio al conjunto de las Comunidades Autónomas y a un número amplio de años. También lleva a cabo cálculos similares a los del margen de maniobra Barea (1998), que estima los gastos no financieros de los Presupuestos Generales del Estado que no provienen de decisiones adoptadas en ejercicios anteriores.

[2] Únicamente vamos a exponer las líneas básicas del proceso de clasificación. Para un mayor detalle, por capítulos, artículos y, en algunos casos, conceptos presupuestarios, véase Barberán (2003).

pio carácter del gasto —definido por la clasificación económica— y de su finalidad —definida por la clasificación por programas—. La perspectiva temporal básica que adoptamos es la del corto plazo.

2.1. **Clasificación de los recursos**

La perspectiva más tradicional en el estudio de los recursos de las comunidades autónomas españolas es la que atiende a la clasificación de los mismos en tres bloques, según que su finalidad principal sea asegurar la financiación de los servicios traspasados, hacer efectivo el principio de solidaridad regional o permitir el ejercicio de la autonomía financiera. Se trata, sin embargo, de una clasificación que ofrece pocas posibilidades en el marco de este trabajo, ya que no discrimina bien entre distintos ingresos según su grado de discrecionalidad.

La tipología que utiliza la Dirección General de Fondos Comunitarios y Financiación Territorial en su Informe anual sobre la financiación de las comunidades autónomas para clasificar los recursos de éstas se adapta notablemente al fin aquí propuesto, ya que atiende, entre otras variables, a la discrecionalidad en su uso y los divide en incondicionados y condicionados. Los recursos clasificados como incondicionados son, efectivamente, recursos de los que la administración autonómica puede disponer con total libertad (discrecionalidad plena), en tanto que los clasificados como condicionados podrían subdividirse en varios subgrupos según el rigor de la condición.

Partiendo de esa tipología procedemos a la reclasificación de los recursos en los cuatro bloques de discrecionalidad antes señalados.

En el bloque de discrecionalidad plena incluimos una amplia variedad de recursos sobre los que la hacienda regional tiene plena capacidad de decisión en cuanto a su uso sin que tenga que dar cuenta de ello a ningún nivel superior de gobierno. Comprende tanto recursos propios como transferidos (transferencias generales), con la característica común de que su uso no está sometido a condición alguna.

En el bloque de discrecionalidad alta incluimos aquellos recursos sobre los que pesa una restricción de carácter general a su libertad de uso. De nuevo coexisten recursos propios y transferidos (transferencias condicionadas en bloque) que tienen en común el encontrarse sometidos a alguna condición de uso que puede calificarse de amplia, como, por ejemplo, su destino exclusivo a un capítulo o a una función de gasto.

En el bloque de discrecionalidad baja incluimos únicamente recursos transferidos (transferencias condicionadas categóricas) cuyo empleo queda sujeto a una restricción estricta fijada por la administración que concede la financiación (central o europea). Esta restricción suele traducirse en términos de proyectos de inversión, programas de transferencias o provisión de servicios perfectamente identificados, seleccionados de mutuo acuerdo entre las administraciones implicadas pero con sujeción a los criterios prefijados por la administración financiadora. Así, una vez establecido el acuerdo en forma de plan, programa, convenio o contrato, la capacidad del gobierno regional para proceder a la reasignación

de créditos, queda limitada a su fase de ejecución, y será más o menos amplia dependiendo de cómo estén definidos los términos de los proyectos o programas.

En el bloque de discrecionalidad nula incluimos sólo recursos transferidos (transferencias condicionadas categóricas) sujetos a una restricción extrema fijada por la administración financiadora: su destino es la concesión de transferencias a otros agentes económicos —públicos o privados— con arreglo a los criterios establecidos por la administración que concede la financiación y sin que quepa ningún margen de maniobra para el gobierno regional, ni siquiera en el proceso de ejecución. Al respecto puede hablarse de meras transferencias gestionadas, cuyo paso por el presupuesto autonómico no tiene ningún efecto, por lo que resultaría perfectamente prescindible y sustituible por una relación directa entre la Administración central o europea y los destinatarios últimos de las ayudas.

El resumen de la clasificación obtenida tras el análisis pormenorizado de los distintos conceptos de ingreso se recoge en el cuadro n.º1.

Cuadro n.º 1

Clasificación de los recursos de las CC AA según el grado de discrecionalidad que permiten en las decisiones presupuestarias

1. Discrecionalidad plena	— Tributos propios. — Recargos sobre tributos estatales. — Tributos cedidos. — Rendimiento de la tarifa autonómica del IRPF. — Tasas afectas a los servicios transferidos. — Precios públicos y otras tasas. — Reintegros y otros ingresos (capítulo 3). — Participación en ingresos generales del Estado. — Participación en ingresos territoriales del Estado por IRPF. — Transferencias por aplicación del Fondo de Garantía. — Participación de las Comunidades Autónomas uniprovinciales y Ciudades Autónomas por recursos provinciales. — Transferencias por coste efectivo de los servicios traspasados. — Ingresos patrimoniales (capítulo 5). — Enajenación de inversiones reales, excepto terrenos ISVA (capítulo 6). — Activos financieros (capítulo 8)
2. Discrecionalidad alta	— Transferencias para la financiación de los servicios traspasados por la Seguridad Social (servicios sanitarios y sociales). — Fondo de Compensación Interterritorial. — Otras transferencias en bloque. — Pasivos financieros (capítulo 9).

Sumario

3. Discrecionalidad baja	— Venta de terrenos ISVA. — Ayudas con cargo al presupuesto general de la Unión Europea (excepto subvenciones del FEOGA-Garantía y otras transferencias gestionadas) (*). — ngresos propios o transferidos destinados a la cofinanciación de las ayudas con cargo al presupuesto general de la Unión Europea (*). — Convenios. — Otras transferencias categóricas.
4. Discrecionalidad nula	— Participación de los entes locales en los ingresos del Estado. — Subvenciones del FEOGA-Garantía. — Contratos programa. — Subvenciones gestionadas. — Otras transferencias gestionadas.

(*) En la perspectiva de largo plazo estos recursos pasarían a clasificarse en el bloque de discrecionalidad alta.

Fuente: Elaboración propia

2.2. **Clasificación de los empleos**

Una vez que los recursos han sido clasificados según los distintos grados de discrecionalidad que permiten las restricciones impuestas de origen debe efectuarse una nueva ronda clasificatoria atendiendo las limitaciones derivadas del empleo de los mismos. El punto de partida para la clasificación de los empleos es el resultado de la clasificación de los recursos. Así, los empleos financiados mediante recursos de discrecionalidad nula deben ser clasificados de igual modo, ya que habrán de ajustarse forzosamente a la finalidad preestablecida. En cambio, entre los restantes empleos caben situaciones diversas: los financiados con ingresos de discrecionalidad baja pueden ser como máximo de discrecionalidad baja, pero también de discrecionalidad nula; los financiados con ingresos de discrecionalidad alta pueden ser tanto de discrecionalidad alta, como baja o nula; y los financiados con ingresos de discrecionalidad plena pueden encontrarse en cualquiera de las cuatro situaciones (discrecionalidad plena, alta, baja o nula)[3].

La posible pérdida de discrecionalidad en las decisiones de gasto, más allá de lo obligado por las condiciones impuestas por el lado del ingreso, tiene que ver con lo que parece una tendencia natural del gasto público que se realiza en un momento dado a consolidarse en el tiempo, haciendo muy difícil su reducción o supresión, cuando menos a corto plazo. Se trata de un fenómeno de solidificación del flujo de gasto según el cual el punto de partida en la elaboración del presupuesto de un ejercicio tiende a ser el gasto realizado en el ejercicio anterior, lo que en la literatura hacendística es conocido como incrementalismo presupuestario.

[3] Esta identificación de los gastos según las condiciones que les impone su fuente de financiación es clave para la estimación final del margen de maniobra presupuestario. En caso de que gastos financiados con ingresos sujetos a condición de uso no se identifiquen y queden clasificados en un bloque de discrecionalidad superior, el margen de maniobra estimado quedará sesgado al alza.

Este fenómeno de solidificación del gasto tiene su origen en factores tanto de demanda como de oferta. Los primeros tienen que ver con la resistencia social a prescindir de los servicios que se estaban disfrutando previamente. Los segundos nacen de las rigideces institucionales (introducidas por normas legales, propias o de gobiernos de nivel superior, o por acuerdos rubricados con terceros) que impiden suprimir o ajustar a la baja el nivel de ciertos gastos y de las inercias del procedimiento de elaboración del presupuesto (ocasionadas, entre otras razones, por las limitaciones técnicas que impiden partir anualmente de cero y por los intereses de los responsables de las unidades administrativas implicadas).

Ahora bien, a efectos del cálculo del margen de maniobra presupuestario, creemos que la atención debe centrarse especialmente en los factores de oferta derivados de las restricciones institucionales, particularmente en las restricciones derivadas de normas legales (distintas de la propia Ley de Presupuestos) o de acuerdos formales suscritos con terceros[4]. La razón es que son factores objetivos sobre los que puede haber consenso para su identificación, lo que no sucede en el caso de las inercias administrativas y las demandas sociales.

Teniendo en cuenta estos criterios clasificamos los empleos en los mismos cuatro bloques de discrecionalidad que los ingresos.

En el bloque de discrecionalidad plena incluimos aquellos empleos que son fruto de decisiones plenamente autónomas del gobierno regional adoptadas en el curso del proceso presupuestario del ejercicio que se analiza. Por su propia naturaleza encajan de modo especial en este bloque las operaciones de capital financiadas con recursos propios.

En el bloque de discrecionalidad alta incluimos los empleos financiados con recursos de discrecionalidad alta y aquellos otros que, siendo financiados con recursos de discrecionalidad plena, se ven afectados por una restricción institucional de carácter general que limita moderadamente la libertad de asignación del gobierno regional dentro del proceso presupuestario del ejercicio analizado. De nuevo, el protagonismo corresponde a las operaciones de capital, financiadas ahora mediante endeudamiento o mediante transferencias condicionadas en bloque.

En el bloque de discrecionalidad baja incluimos los empleos financiados con recursos de discrecionalidad baja y aquellos otros financiados con recursos que permiten mayor margen de discrecionalidad pero que se ven afectados por una restricción institucional que limita sustancialmente la libertad de asignación del gobierno regional. Con ello, la discrecionalidad en la fijación del volumen de gasto o en su destino se reduce a un muy estrecho margen en el proceso de elaboración del presupuesto y en el proceso de ejecución. Los

[4] Dentro de este enfoque, las restricciones que provienen del lado del ingreso encajan perfectamente como restricciones legales. Del resto de restricciones legales, hay que tener presente que las dadas por niveles de gobierno superiores resultan inamovibles para el gobierno regional, mientras que las adoptadas por el propio gobierno regional pueden ser modificadas por éste dentro de ciertos márgenes, lo que influye, sin duda, en su margen de maniobra; en el caso de los acuerdos suscritos con terceros, la clave está en las condiciones pactadas para su rescisión.

gastos de personal y funcionamiento constituyen su componente más característico.

En el bloque de discrecionalidad nula incluimos los empleos financiados con recursos de discrecionalidad nula junto a cualesquiera otros afectados por una restricción institucional extrema: no existe ningún margen de discrecionalidad a corto plazo para el gobierno regional, ni en el volumen, ni en el destino del gasto, tanto en el proceso de aprobación como en el de ejecución del presupuesto. Cumplen estas características las trasferencias gestionadas, los gastos asociados al servicio de la deuda y los compromisos plurianuales de gasto.

El resumen de la clasificación obtenida tras el análisis pormenorizado de las distintas partidas de gasto puede verse en el cuadro n.º 2.

Cuadro n.º 2

Clasificación de los empleos de las CC AA según el grado de discrecionalidad que permiten en las decisiones presupuestarias

1. Discrecionalidad plena	— Gastos diversos, procesos electorales, estudios y trabajos técnicos, servicios nuevos y otros gastos (Capítulo 2) — Transferencias corrientes a empresas privadas y a familias e instituciones sin fines de lucro (excepto por enseñanza concertada) financiadas con recursos propios. — Inversiones reales financiadas con recursos propios. — Transferencias de capital financiadas con recursos propios (excepto a la Administración del Estado y a Corporaciones locales). — Activos financieros.
2. Discrecionalidad alta	— Gastos de altos cargos, personal eventual de gabinetes, personal laboral eventual, otro personal e incentivos al rendimiento. — Material y suministros (proporcional a gastos de personal) (Capítulo 2) — Transferencias corrientes a Corporaciones locales financiadas con recursos propios. — Inversiones reales financiadas con endeudamiento. — Transferencias de capital financiadas con endeudamiento. — Transferencias de capital a Corporaciones locales financiadas con recursos propios.
3. Discrecionalidad baja	— Gastos de personal funcionario y personal laboral fijo. — Reparaciones y conservación (Capítulo 2). — Material y suministros (proporcional a gastos de personal) (Capítulo 2). — Transferencias corrientes cofinanciadas por la Unión Europea (*). — Transferencias corrientes a la Administración del Estado, a Organismos públicos y a familias e instituciones sin fines de lucro (por enseñanza concertada) financiadas con recursos propios (*). — Inversiones reales financiadas con recursos finalistas de la Administración central. — Inversiones reales cofinanciadas por la Unión Europea (*). — Transferencias de capital cofinanciadas por la Unión Europea (*). — Transferencias de capital a la Administración del Estado financiadas con recursos propios (*).

Sumario

4. Discrecionalidad nula	— Arrendamientos (Capítulo 2) (*). — Gastos financieros (Capítulo 3). — Transferencias corrientes financiadas con recursos finalistas de la Administración central o de la Unión Europea (incluidas las financiadas por el FEOGA-Garantía en el marco de la PAC). — Transferencias de capital financiadas con recursos finalistas de la Administración central o de la Unión Europea. — Compromisos de gastos plurianuales adoptados en ejercicios anteriores (Capítulos 2, 4, 6, 7 y 8 del presupuesto) (**). — Pasivos financieros.

(*) En la perspectiva a largo plazo estos empleos pasan a clasificarse en el bloque de discrecionalidad alta.
(**) En la perspectiva a largo plazo estos empleos se reclasifican, según su naturaleza, en el bloque de discrecionalidad plena, alta o baja.

Fuente: Elaboración propia

2.3. **Medición del margen de maniobra**

Una vez clasificados los distintos tipos de recursos y empleos en cuatro bloques, se dispone de la herramienta necesaria para medir el margen de maniobra presupuestario. Para que esta herramienta resulte operativa, habrá que seleccionar una fuente adecuada de información sobre los ingresos y gastos del gobierno regional[5], procediéndose a continuación a traspasar los datos de la clasificación presupuestaria de origen a esta nueva clasificación. Tras este proceso de cálculo tendremos todos los ingresos y gastos distribuidos entre los cuatro bloques y, por tanto, ya podrá analizarse el grado de discrecionalidad con el que ha contado el gobierno regional en la gestión de su hacienda.

Al objeto de disponer de un indicador que sintetice en una única cifra (un porcentaje) el concepto de margen de maniobra, de modo que sea más fácil su manejo, reduciremos los cuatro bloques de discrecionalidad que se han utilizado hasta ahora a sólo dos: 1) discrecionalidad alta o plena; 2) discrecionalidad baja o nula. Con ello se establece la distinción básica entre lo que son recursos mediante los cuales el gobierno regional puede llevar a la práctica su propia política económica —en sentido amplio— y aquellos a través de los cuales son otros gobiernos los que aplican su política convirtiendo al regional en mero ejecutor. Del mismo modo, esto supone distinguir entre los empleos discrecionales a través de los cuales el gobierno regional puede, en el corto plazo, introducir cambios en su política presupuestaria y aquellos otros en los que se ve tan fuertemente condicionado que se limita a hacer gestión o administración de servicios y transferencias.

Esta clasificación dicotómica de los recursos nos permite obtener una ratio representativa del «margen de maniobra

[5] Dado que la noción de discrecionalidad que estamos utilizando se refiere a la capacidad del gobierno para tomar decisiones de asignación de fondos tanto en la fase de elaboración y aprobación del presupuesto como en la de ejecución, no debería tomarse como fuente de datos el presupuesto sino la liquidación del mismo.

permitido por los ingresos»en tanto que la clasificación de los empleos nos permite obtener la ratio representativa del «margen de maniobra realmente utilizado». Sin embargo, el margen de maniobra del que verdaderamente disfruta el gobierno regional en un ejercicio presupuestario no queda totalmente explicado por las cifras de ingresos y gastos ejecutados en ese ejercicio. Así, introduciendo algunos ajustes a esas cifras pueden conseguirse nuevos valores que proporcionan una visión más completa del margen de maniobra, en concreto pueden obtenerse medidas de los márgenes de maniobra potencial y neto.

El «margen de maniobra potencial» lo definimos como el margen que podría haber utilizado el gobierno regional si hubiese hecho uso de todos los recursos no tributarios que tiene a su disposición[6]. Para su cálculo sumamos a los empleos de discrecionalidad alta o plena los recursos potenciales de naturaleza financiera que operan a modo de reservas de ese gobierno: el remanente de tesorería acumulado al final del ejercicio[7] y el margen de endeudamiento permitido por los escenarios de consolidación presupuestaria vigentes y no utilizado al término del ejercicio[8]. Este margen potencial marca el límite al que en cada ejercicio podría haber llegado el gobierno regional si hubiese decidido introducir un gran cambio puntual en su política presupuestaria. Señala una pauta de comportamiento presupuestario claramente excepcional y solo sostenible en el muy corto plazo.

El «margen de maniobra neto» lo definimos como el margen que como máximo podría haber alcanzado el gobierno regional si no hubiese hecho uso de fuentes extraordinarias de financiación. Para su cálculo restamos a los empleos de dicrecionalidad alta o plena los recursos extraordinarios que han sido utilizados por ese gobierno, en particular: el incremento neto de pasivos financieros[9], la disminución

[6] No tomamos en consideración el margen potencial que surge de la posibilidad de que disponen los gobiernos regionales de incrementar los ingresos de tipo tributario mediante la adopción de las medidas discrecionales (incremento de tipos impositivos, imposición de recargos, etc.) en el marco de la legislación actual sobre su sistema de financiación. Y ello, no porque el margen de maniobra que surge por esa vía sea irrelevante, sino porque su naturaleza es tal que no resulta factible ni sería riguroso el cálculo de la recaudación máxima posible, con lo que tampoco lo es el cálculo de los ingresos potenciales que se están dejando de obtener en cada momento.

[7] El remanente de tesorería existente al término de un ejercicio presupuestario o, lo que es lo mismo, los activos líquidos netos que quedan a disposición del gobierno para su uso en la financiación de gastos de ejercicios posteriores, es una fuente del margen de maniobra, por lo que su existencia indica que el margen de maniobra realmente utilizado no ha alcanzado su límite potencial (se adopta el supuesto simplificador de que el remanente puede utilizarse con un amplio grado de discrecionalidad por el gobierno regional y, por tanto, se obvian las restricciones que habitualmente impone la legislación para su uso).

[8] La aprobación de los escenarios de consolidación presupuestaria a partir de 1991, en los que se establecen límites obligatorios para cada una de las comunidades autónomas en lo referente a la magnitud del déficit (necesidad de financiación) y de la deuda viva, permite introducir en el análisis del margen de maniobra potencial el margen de endeudamiento no utilizado en cada ejercicio. De tal modo que si no se ha agotado el límite de endeudamiento fijado quiere decir que el margen de maniobra realmente utilizado no ha alcanzado su límite potencial (se adopta el supuesto simplificador de que los ingresos por pasivos financieros se aplicarían íntegramente al bloque de gastos de discrecionalidad alta). Obviamente, el cambio de marco de disciplina presupuestaria que se produjo en 2003, con la entrada en vigor de la Ley General de Estabilidad Presupuestaria, ha supuesto la desaparición de esta fuente de margen de maniobra.

[9] Si la variación neta de pasivos financieros resulta negativa no será tomada en cuenta, ya que en tal

neta de activos financieros[10], la enajenación de inversiones reales[11] y la variación neta del remanente de tesorería[12]. Este margen neto indica la senda de comportamiento presupuestario que es perfectamente sostenible a largo plazo y anticipa, por tanto, la senda que inevitablemente acabará siguiendo el margen realmente utilizado, salvo que concurran circunstancias extraordinarias como cambios en el sistema de financiación o en los servicios traspasados a las comunidades autónomas.

3. RESULTADOS DE LA APLICACIÓN A LA COMUNIDAD AUTÓNOMA DE ARAGÓN

El periodo sometido a estudio abarca desde 1986, último año de vigencia del llamado «periodo transitorio» del sistema de financiación de las comunidades autónomas —momento en que la joven administración autonómica puede darse por consolidada—, hasta 2000. Los datos analizados provienen de la Cuenta General que anualmente elabora la Intervención General del Gobierno de Aragón y se refieren a derechos y obligaciones reconocidos.

3.1. El margen de maniobra permitido por los ingresos

De la aplicación de los criterios de clasificación a las liquidaciones del presupuesto de ingresos se obtienen los resultados incluidos en el cuadro n.º 3.

La evolución de los cuatro bloques de discrecionalidad permite dividir el periodo estudiado en dos etapas claramente diferenciadas. La primera hasta 1993, en la que el margen de maniobra se mantiene entorno al 85% del total de ingresos. La segunda, desde 1994 hasta el final del periodo de estudio, marcada por una intensa reducción de ese margen (26 puntos porcentuales), aunque en los dos últimos años se produce una notable recuperación.

Esta evolución muestra una especial sensibilidad a los cambios institucionales, particularmente los operados en el sistema de financiación y en la distribución de competencias entre Estado y Comunidad Autónoma. Entre esos cambios cabe señalar los siguientes: la reforma

caso los pasivos no han sido fuente de margen de discrecionalidad sino que, de hecho, lo han reducido.

[10] Si la variación neta de activos financieros da un valor positivo tampoco será tomada en cuenta, ya que en lugar de contribuir a ampliar el margen de discrecionalidad está consumiéndolo.

[11] Estos ingresos serán tenidos en cuenta sólo en los casos en que su uso no se encuentre condicionado a la inversión en el mismo tipo de bienes que se enajenaron.

[12] La variación del remanente de tesorería cubre los desajustes entre derechos reconocidos y obligaciones contraídas en un ejercicio (diferencia entre ingresos y gastos totales finalmente ejecutados): una disminución del remanente se da cuando el saldo ha sido negativo, y un incremento, cuando el saldo ha sido positivo. Por ello, la disminución del remanente de tesorería tendrá el mismo tratamiento que el incremento de pasivos, ya que para la obtención del margen de maniobra realmente utilizado ha sido preciso recurrir no sólo a los pasivos sino a la incorporación de remanentes por la cuantía del déficit; mientras que el incremento del remanente se restará de los derechos reconocidos por incremento de pasivos, ya que puede suponerse que parte de la deuda contraída se ha destinado a la financiación del superávit y, por tanto, no resultaba precisa para la obtención del margen de maniobra realmente utilizado. Cuando el incremento del remanente de tesorería sea mayor que la suma de incremento neto de pasivos financieros, disminución neta de activos financieros y enajenación de inversiones reales, deberá darse un valor cero al conjunto, con lo que el margen neto será igual al margen realmente utilizado.

Cuadro n.º 3

Evolución del grado de discrecionalidad de los recursos de la Comunidad Autónoma de Aragón, 1986-2000.

Derechos reconocidos

	Discrecionalidad plena			*Discrecionalidad alta*			*Discrecionalidad baja*			*Discrecionalidad nula*			*Total ingresos*		
	millones ptas	% total ingr.	% VABcf	millones ptas	% total ingr.	% VABcf	millones ptas	% total ingr.	% VABcf	millones ptas	% total ingr.	% VABcf	millones ptas	% total ingr.	% VABcf
1986	20.703,2	61,8	2,0	5.087,5	15,2	0,5	6.039,6	18,0	0,6	1.694,2	5,1	0,2	33.524,5	100,0	3,2
1987	27.020,3	68,9	2,4	4.465,7	11,4	0,4	4.450,5	11,3	0,4	3.299,1	8,4	0,3	39.235,7	100,0	3,5
1988	32.477,6	74,9	2,5	4.421,3	10,2	0,3	3.468,6	8,0	0,3	3.019,7	7,0	0,2	43.387,1	100,0	3,3
1989	39.647,5	66,5	2,7	10.910,0	18,3	0,8	5.647,7	9,5	0,4	3.446,6	5,8	0,2	59.651,8	100,0	4,1
1990	43.138,7	53,9	2,7	27.741,7	34,7	1,7	4.951,6	6,2	0,3	4.214,3	5,3	0,3	80.046,3	100,0	5,0
1991	44.727,0	51,0	2,6	30.830,5	35,1	1,8	8.855,4	10,1	0,5	3.310,5	3,8	0,2	87.723,3	100,0	5,0
1992	54.471,2	66,1	2,9	10.688,8	13,0	0,6	14.770,0	17,9	0,8	2.503,0	3,0	0,1	82.433,0	100,0	4,5
1993	52.599,0	57,2	2,7	21.921,0	23,8	1,1	14.593,1	15,9	0,8	2.882,9	3,1	0,1	91.996,0	100,0	4,8
1994	59.451,9	41,6	2,9	33.019,3	23,1	1,6	11.825,9	8,3	0,6	38.481,9	27,0	1,9	142.778,9	100,0	7,0
1995	52.647,5	41,2	2,4	0,0	0,0	0,0	16.771,4	13,1	0,8	58.330,3	45,7	2,7	127.749,2	100,0	5,8
1996	77.130,5	42,5	3,3	22.670,9	12,5	1,0	22.119,7	12,2	1,0	59.414,3	32,8	2,6	181.335,4	100,0	7,8
1997	87.171,5	44,4	3,6	26.046,5	13,3	1,1	22.510,6	11,5	0,9	60.635,2	30,9	2,5	196.363,8	100,0	8,0
1998	93.566,8	43,2	3,7	17.162,0	7,9	0,7	38.745,6	17,9	1,5	67.027,6	31,0	2,6	216.502,0	100,0	8,5
1999	178.419,4	59,5	6,7	16.675,4	5,6	0,6	33.939,7	11,3	1,3	70.800,8	23,6	2,7	299.835,3	100,0	11,3
2000	187.524,0	59,6	6,6	16.166,4	5,1	0,6	28.865,6	9,2	1,0	82.002,4	26,1	2,9	314.558,4	100,0	11,1

Fuente: Elaboración propia

Sumario

del sistema de financiación de las comunidades autónomas con la cesión parcial del IRPF, en 1997 y 1998; las reformas del FCI , con efecto sobre todo en 1987 y 1992; el traspaso de los servicios del Imserso, en 1997, y el traspaso de los servicios de educación desde 1999. También influyen significativamente en esta evolución la política de endeudamiento seguida por el gobierno regional y ciertas decisiones de naturaleza no económica sino contable. Las decisiones más destacadas a este respecto son contabilizar como ingreso las transferencias gestionadas provenientes del FEOGA-Garantía en el marco de la PAC[14], que tiene efecto desde 1994, y la modificación, en 1995, de los criterios de contabilización de los derechos reconocidos por pasivos financieros.

Los ingresos que pueden utilizarse con total discrecionalidad por parte del gobierno regional en 2000 representan aproximadamente el 60% del total, y si a éstos se suman los ingresos que pueden utilizarse con un grado alto de discrecionalidad (algo más del 5%), se obtiene como resultado un margen de maniobra equivalente al 65% del total de los ingresos de la hacienda aragonesa[15].

En términos del VAB de Aragón se obtiene, como no puede ser de otro modo en un contexto de crecimiento de los ingresos de la Comunidad Autónoma, una senda claramente creciente de todos los bloques. Aunque la evolución del bloque de discrecionalidad alta presenta altibajos provocados por las decisiones del gobierno regional sobre el recurso al crédito.

3.2. El margen de maniobra realmente utilizado: la discrecionalidad en el gasto

La aplicación de los criterios de clasificación a las liquidaciones del presupuesto de gastos de la Comunidad Autónoma de Aragón conduce a los resultados incluidos en el cuadro n° 4.

Como en el caso de los ingresos, el análisis de la evolución conjunta de los cuatro bloques de discrecionalidad permite dividir el periodo estudiado en dos etapas claramente diferenciadas, siendo de nuevo el año 1993 el último de la primera.

En esta primera etapa, el bloque de discrecionalidad plena se mantiene en torno al 28% como media, mientras el bloque de discrecionalidad alta —con mayores variaciones interanuales— se vendría a situar en torno al 20%. En cualquier caso, la combinación de ambos bloques da lugar a un margen de maniobra que no cae en ningún momento por debajo del 45% y que en algunos años llega hasta el 60%.

En la segunda etapa, el protagonismo lo adquiere el bloque de discrecionalidad nula, aunque compartido con el de discrecionalidad baja, lo que deja reducido el margen de maniobra a una banda de entre el 15% y el 20% (17,7%, media del período).

De nuevo, la evolución temporal de estos márgenes de maniobra se ha visto condicionada fuertemente por el traspaso de nuevos servicios del Estado a la co-

[14] Constituye un claro ejemplo de lo que ha dado en llamarse «contabilidad creativa».

[15] El fuerte peso que alcanzan lo bloques de discrecionalidad baja y nula se reduce drásticamente cuando se excluyen las transferencias del FEOGA-Garantía.

Cuadro n.º 4

Evolución del grado de discrecionalidad de los recursos de la Comunidad Autónoma de Aragón, 1986-2000.

Obligaciones reconocidas

	Discrecionalidad plena			*Discrecionalidad alta*			*Discrecionalidad baja*			*Discrecionalidad nula*			*Total gastos*		
	millones ptas	% total gtos.	% VABcf	millones ptas	% total gtos.	% VABcf	millones ptas	% total gtos.	% VABcf	millones ptas	% total gtos.	% VABcf	millones ptas	% total ingr.	% VABcf
1986	9.569,6	33,2	0,9	4.566,5	15,9	0,4	12.734,3	44,2	1,2	1.922,6	6,7	0,2	28.793,0	100,0	2,8
1987	9.445,4	27,7	0,8	6.009,7	17,6	0,5	14.822,0	43,4	1,3	3.867,7	11,3	0,3	34.144,8	100,0	3,0
1988	12.225,2	32,8	0,9	4.194,8	11,3	0,3	17.506,2	47,0	1,3	3.322,0	8,9	0,3	37.248,1	100,0	2,9
1989	12.335,0	25,9	0,9	9.854,4	20,7	0,7	20.819,7	43,7	1,4	4.654,2	9,8	0,3	47.663,3	100,0	3,3
1990	16.472,7	26,1	1,0	18.514,4	29,3	1,2	22.334,4	35,4	1,4	5.841,1	9,2	0,4	63.162,6	100,0	4,0
1991	21.363,8	25,3	1,2	29.755,2	35,3	1,7	27.346,1	32,4	1,6	5.876,9	7,0	0,3	84.342,0	100,0	4,8
1992	31.325,5	32,5	1,7	22.524,8	23,4	1,2	34.866,4	36,2	1,9	7.693,0	8,0	0,4	96.409,7	100,0	5,2
1993	21.937,6	22,9	1,1	21.253,5	22,2	1,1	38.509,4	40,3	2,0	13.965,2	14,6	0,7	95.665,6	100,0	4,9
1994	25.736,9	18,0	1,3	30.624,0	21,5	1,5	36.509,6	25,6	1,8	49.740,4	34,9	2,4	142.611,0	100,0	7,0
1995	9.773,2	7,1	0,4	7.140,5	5,2	0,3	45.856,2	33,2	2,1	75.500,9	54,6	3,4	138.270,8	100,0	6,3
1996	9.109,8	5,1	0,4	25.730,3	14,4	1,1	57.336,2	32,0	2,5	86.969,4	48,5	3,7	179.145,7	100,0	7,7
1997	13.280,1	6,9	0,5	24.962,0	12,9	1,0	56.563,1	29,3	2,3	98.344,2	50,9	4,0	193.149,5	100,0	7,9
1998	14.414,3	6,9	0,6	23.986,4	11,4	0,9	67.684,3	32,2	2,7	104.171,5	49,5	4,1	210.256,5	100,0	8,3
1999	21.930,4	7,5	0,8	22.393,4	7,7	0,8	134.033,1	46,1	5,0	112.651,0	38,7	4,2	291.007,9	100,0	11,0
2000	32.167,9	9,8	1,1	23.634,8	7,2	0,8	146.097,0	44,5	5,2	126.661,4	38,6	4,5	328.561,1	100,0	11,6

Fuente: Elaboración propia

Sumario

munidad autónoma y por la financiación asociada a los mismos. También han tenido una repercusión indiscutible las sucesivas reformas del sistema de financiación de las comunidades autónomas de régimen común y algunas decisiones presupuestarias tomadas autónomamente por el gobierno regional, como la de recurrir en mayor o menor medida al endeudamiento o las más formales de contabilizar las transferencias de la PAC y de modificar los criterios de contabilización de los derechos reconocidos por pasivos financieros[16].

El análisis en términos del VAB de Aragón evidencia la distinta evolución seguida por los bloques de discrecionalidad alta y plena y por los de baja y nula. Los primeros mantienen un peso apenas alterado a lo largo de los quince años estudiados, casi insensible a las fuertes transformaciones sufridas por el sector público autonómico de Aragón[17], en tanto los segundos crecen casi al mismo ritmo que lo hace el tamaño del presupuesto regional.

3.3. **Otras medidas del margen de maniobra: margen potencial y margen neto**

Conforme a la definición de las medidas de los márgenes de maniobra potencial y neto hecha en el apartado anterior y tras su aplicación a las liquidaciones del presupuesto de la Comunidad Autónoma de Aragón en el periodo 1986 a 2000, se obtienen los resultados que se presentan en el cuadro nº 5.

El análisis de estos resultados permite identificar dos fenómenos importantes. En primer lugar, el proceso de convergencia de los márgenes potencial y neto con el realmente utilizado, que se inicia a comienzos de los años noventa. En segundo lugar, la tendencia decreciente de los márgenes de maniobra medidos en términos del total de las obligaciones reconocidas del presupuesto, a la vez que la tendencia casi horizontal cuando se miden en términos del VAB regional. Esta evolución se ilustra con claridad en los gráficos nºs 1 y 2.

Esto pone de manifiesto que el crecimiento extraordinario del tamaño del presupuesto de la Comunidad Autónoma a lo largo del periodo de estudio no se ha traducido en ganancias reales en el margen de discrecionalidad, ya que sobre todo están creciendo los gastos que permiten un bajo o nulo margen de maniobra. Además, la convergencia de los márgenes real y potencial ha traído como consecuencia principal que sean cada vez más reducidas las posibilidades de variación del margen de maniobra a corto plazo, salvo que se produzcan nuevos traspasos de servicios ligados a competencias asumidas por la Comunidad Autónoma o que ésta pueda acceder a nuevas fuentes de financiación distintas de las estrictamente financieras[18].

[16] Nota explicativa: «contabilidad creativa», a desarrollar por el autor.

[17] En los años 1990 a 1994 se observa un mayor peso de los bloques de discrecionalidad alta y plena motivado por la expansión del endeudamiento de la Comunidad Autónoma.

[18] Así habrá ocurrido en 2002 con el traspaso de los servicios de asistencia sanitaria y, en cierto modo, con la aplicación del nuevo sistema de financiación de las comunidades autónomas de régimen común.

Cuadro n.º 5

Evolución del margen de maniobra de la Comunidad Autónoma de Aragón, 1986-2000.

Obligaciones reconocidas

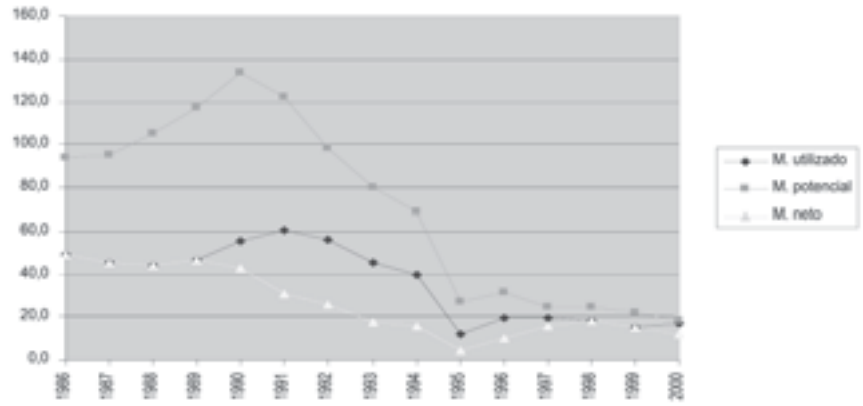
	Margen de maniobra utilizado			*Margen de maniobra potencial*			*Margen de maniobra neto*		
	millones ptas	% total gtos.	% VABcf	millones ptas	% total gtos.	% VABcf	millones ptas	% total gtos.	% VABcf
1986	14.136,1	49,1	1,4	26.964,7	93,7	2,6	14.136,1	49,1	1,4
1987	15.455,1	45,3	1,4	32.496,0	95,2	2,9	15.455,1	45,3	1,4
1988	16.420,0	44,1	1,3	39.228,4	105,3	3,0	16.420,0	44,1	1,3
1989	22.189,4	46,6	1,5	55.847,2	117,2	3,9	22.189,4	46,6	1,5
1990	34.987,1	55,4	2,2	84.287,0	133,4	5,3	27.009,7	42,8	1,7
1991	51.119,0	60,6	2,9	103.122,6	122,3	5,9	25.866,6	30,7	1,5
1992	53.850,3	55,9	2,9	94.495,6	98,0	5,1	25.010,5	25,9	1,4
1993	43.191,0	45,1	2,2	76.266,9	79,7	3,9	16.584,3	17,3	0,9
1994	56.361,0	39,5	2,8	98.116,8	68,8	4,8	22.419,2	15,7	1,1
1995	16.913,7	12,2	0,8	37.278,7	27,0	1,7	6.372,8	4,6	0,3
1996	34.840,1	19,4	1,5	56.343,3	31,5	2,4	17.930,0	10,0	0,8
1997	38.242,2	19,8	1,6	46.862,9	24,3	1,9	29.972,7	15,5	1,2
1998	38.400,7	18,3	1,5	51.243,9	24,4	2,0	38.400,7	18,3	1,5
1999	44.323,8	15,2	1,7	63.340,2	21,8	2,4	44.323,8	15,2	1,7
2000	55.802,7	17,0	2,0	60.176,5	18,3	2,1	39.235,6	11,9	1,4

Fuente: Elaboración propia

Sumario

Gráfico n.º 1

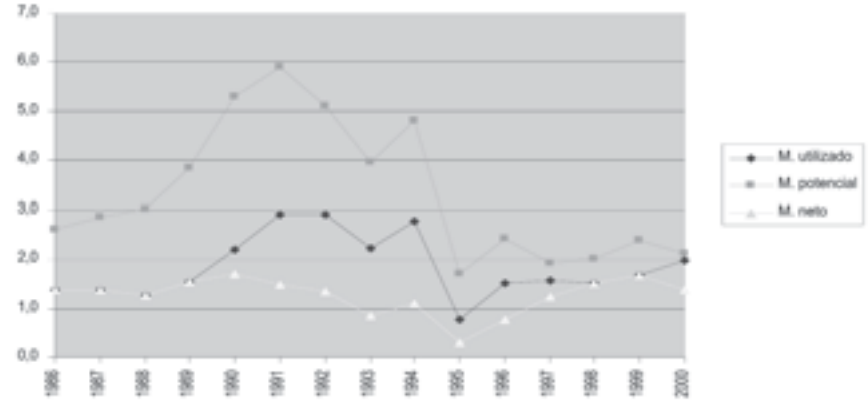
Evolución del margen de maniobra presupuestario
(en % s/total gastos)



Fuente: Elaboración propia

Gráfico n.º 2

Evolución del margen de maniobra presupuestario
(en % s/VAB regional)



Fuente: Elaboración propia

Sumario

4. **CONCLUSIONES**

La capacidad con que cuenta en cada momento un gobierno para decidir de modo discrecional sobre el destino de los recursos presupuestarios que administra afecta directamente a sus posibilidades de variar, a corto plazo, las políticas de gasto que están en marcha y de dar inicio a otras nuevas, con lo que, indirectamente, afecta también a su ortodoxia en la gestión presupuestaria. En particular, cuando el margen de maniobra disponible es insuficiente para hacer efectivas las políticas deseadas por el gobierno, es previsible que aumente la tendencia a la indisciplina presupuestaria.

En este trabajo se ha desarrollado un modelo de análisis que pretende obtener la medida del grado de discrecionalidad o margen de maniobra presupuestario de las comunidades autónomas. La clasificación de los recursos y empleos en cuatro bloques según las posibilidades que ofrecen a los gobiernos regionales para influir en su destino a lo largo del proceso presupuestario constituye el rasgo más característico del método desarrollado. Esta clasificación permite introducir en el análisis una riqueza de matices mayor que la simple división entre empleos discrecionales y no discrecionales, aunque, finalmente, optamos por la sencillez de la clasificación dicotómica a la hora de expresar en una sola cifra (un porcentaje) el margen de maniobra presupuestario.

Además, al objeto de conseguir una mayor profundidad analítica en cuanto a la valoración de este margen de maniobra, hemos complementado la medida que se deduce de considerar la totalidad de los derechos y obligaciones reconocidos en el ejercicio presupuestario (margen de maniobra realmente utilizado) con otras dos medidas del margen de maniobra (margen potencial y margen neto), fruto de la adición o sustracción, según el caso, de ciertas fuentes extraordinarias de financiación.

Del análisis que ha sido necesario realizar para la clasificación de las operaciones presupuestarias se desprende que, en cuanto al origen de los recursos, las restricciones a la libertad de asignación de los gobiernos regionales son ajenas a su propia actuación, ya que vienen impuestas por los gobiernos de nivel superior (Administración central española o Administración de las Comunidades Europeas) a través, principalmente, de la concesión de transferencias condicionadas. En tanto que, si se atiende a los empleos de esos recursos, se observa una combinación de restricciones externas e internas a los gobiernos regionales. Las más características de entre las primeras son las que tienen su origen en la fuente de financiación y en las normas que rigen la vida laboral de los empleados públicos, mientras que las segundas están asociadas a la vigencia de regulaciones y compromisos adoptados por los propios gobiernos regionales en ejercicios pasados.

Los resultados obtenidos, a partir de la aplicación realizada a la Comunidad Autónoma de Aragón, confirman, como era de esperar, una situación extremadamente dispar entre el grado de discrecionalidad de los recursos y el de los empleos. Así, en el ejercicio 2000, el margen de maniobra permitido por los ingresos asciende al 65 % del total de los ingresos de la hacienda aragonesa (7,2 % del VAB regional), en tanto que la discrecionalidad en las decisiones de gasto —lo que hemos llamado margen de maniobra realmente utilizado—

Sumario

asciende al 17 % de los gastos totales del gobierno aragonés (1,9 % del VAB regional). Resultados que indican que el margen de maniobra es bajo en relación con el volumen total del presupuesto, pero de ninguna manera despreciable en términos absolutos o si se pone en relación con el VAB. De modo que el proceso presupuestario no parece ser un trámite irrelevante para las tareas de asignación de recursos y de redistribución.

El análisis de la evolución de los distintos bloques de discrecionalidad pone de manifiesto que las sucesivas reformas del sistema de financiación de las comunidades autónomas de régimen común apenas han incidido en el margen de maniobra presupuestario, a través de la transformación de financiación condicionada en incondicionada y de la aportación de recursos adicionales al sistema. Mientras que el proceso de ampliación de competencias y traspasos de nuevos servicios a las comunidades autónomas (en especial servicios públicos básicos, como la educación), al incrementar el peso del gasto corriente (en particular el gasto de personal) en detrimento del de capital, está teniendo como efecto una fuerte pérdida de margen de maniobra en términos del total de gastos del presupuestos autonómico y, por tanto, una mayor limitación a las posibilidades de introducir cambios significativos en la dimensión financiera de las políticas de los gobiernos autonómicos.

Por tanto, los nuevos traspasos de servicios, habidos con posterioridad al término del periodo de estudio, han debido seguir recortando el margen de maniobra de las comunidades autónomas de bajo nivel de competencias hasta niveles semejantes a los de las comunidades de alto nivel de competencias y régimen común de financiación. En este contexto, la recuperación de cierto margen de maniobra, si no media un incremento de la financiación aportada al sistema, parece depender directamente de que las comunidades autónomas incrementen la presión fiscal en su territorio, ejerciendo las posibilidades que tienen en materia de tributos propios y de tributos cedidos con competencias normativas. Sin embargo, la experiencia reciente indica que los costes electorales de esta vía de incremento del margen de maniobra pueden inducir a las comunidades a seguir otras vías menos ortodoxas de financiación.

REFERENCIAS BIBLIOGRÁFICAS

BARBERÁN, R. (2003): *La medición del grado de discrecionalidad de las decisiones presupuestarias de la Comunidades Autónomas*. Instituto de Estudios Fiscales, Documentos 13/03. Madrid.

BAREA, J. (1998): *Disciplina presupuestaria e integración de España en la Unión Europea*. Real Academia de Ciencias Morales y Políticas. Madrid.

GOBIERNO DE ARAGÓN (1986 a 2000): *El Presupuesto de la Comunidad Autónoma de Aragón para ....*Departamento de Economía y Hacienda. Zaragoza.

GOBIERNO DE ARAGÓN (1987 a 2001): *Cuenta General de la Comunidad Autónoma de Aragón para ....*Departamento de Economía y Hacienda. Inédito.

GÓMEZ J., GARCÍA A., JORRÍN J. A., MARTÍN J. V. Y TRUEBA J. C. (1990): *Gasto comprometido del Estado*. Instituto de Estudios Fiscales, Monografía nº 83. Madrid.

Sumario

Sumario

AUTORES

BALLART, Xavier. Es profesor titular de Ciencias Políticas y de la Administración de la Universidad Autónoma de Barcelona. Licen ciado con grado en Derecho por la Universidad de Barcelona en 1985; Premio de la Escuela de Administración Pública de Cataluña en 1986; Master of Public Affairs por la Universidad Indiana University Bloomington, School of Public and Environmental Affairs Award, 1988; Doctor en Ciencias Poíticas y de la Administración por la Universidad Autónoma de Barcelona en 1991. Sus principales líneas de investigación son la evaluación de los servicios públicos, la administración pública en España, y la innovación en gestión pública.

BARBERÁN ORTÍ, Ramón. Es profesor titular de Economía Aplicada del Departamento de Estructura Económica y Economía Pública de la Universidad de Zaragoza. Licenciado en Ciencias Económicas y Empresariales, y Doctor en Economía. Sus principales líneas de investigación y docencia son la economía aplicada, la economía pública y el federalismo fiscal.

CASTELLANOS MADUELL, Albert. Es profesor asociado del Departamento de Economía y Empresa de la Universidad Pompeu Fabra. Licenciado en Economía por la Universidad Pompeu Fabra en el año 2000. Master en Economía Pública por la Universidad de York (R.U.). Sus principales líneas de investigación y docencia son la hacienda pública, la economía pública y el federalismo fiscal.

Sumario

Cirera León, Ana. Investigadora del Instituto de Desarrollo Regional, Fundación Universitaria. Licenciada en Ciencias Económicas y Empresariales por la Universidad de Sevilla. Curso de postgrado en Economía Regional y Urbana en la Sheffield Hallam University (Reino Unido). Coordinadora en la línea de Evaluación de Políticas Públicas del Instituto de Desarrollo Regional. Ha trabajado como consultor senior en Grant Thornton Consultores (México); analista de inversiones en MedGroup Internacional, División de Soros Real Estate, (Barcelona), asistente de investigación en IESE, Universidad de Navarra (Barcelona), e investigadora del Instituto de Desarrollo Regional. Ha participado en la redacción de la *Guía para la Evaluación de Políticas Públicas*. Desde 2005 es secretaria y tesorera de la Sociedad Española de Evaluación.

Gómez Sancho, José María. Licenciado en Ciencias Económicas Empresariales (sección Economía) y Doctor en Economía por la Universidad de Zaragoza. En la actualidad es Profesor Ayudante de Economía Aplicada en la Facultad de Ciencias Económicas y Empresariales de la Universidad de Zaragoza. Sus principales líneas investigadoras se concretan en el estudio de la eficiencia productiva, principalmente en el sector educativo universitario.

González López-Valcarcel, Beatriz. Es catedrática de Economía Aplicada de la Universidad de Las Palmas de Gran Canaria Pompeu Fabra. Licenciada con Grado en Economía por la Universidad de Santiago de Compostela en 1975, Doctora en Economía por la Universidad de La Laguna en 1983. Sus principales líneas de investigación y docencia son la economía de la salud, los métodos cuantitativos y de simulación, y su aplicación al diseño de las organizaciones sanitarias, y la financiación y gestión de los servicios sanitarios; la economía del tabaco y la economía del medicamento. Entre otras responsabilidades, es Presidenta de la Asociación Española de la Economía de la Salud (AES) y miembro de diferentes consejos asesores y editoriales de revistas de ámbito nacional e internacional.

López Casasnovas, Guillem. Es catedrático de Economía de la Universidad Pompeu Fabra. Licenciado con premio extraordinario en Ciencias Económicas y Empresariales por la Universidad de Barcelona en 1978. Licenciado en Derecho en 1979. Doctor en Economía Pública por la Universidad de York (R.U.). En 1996 fundó el Centro de Investigación en Economía y Salud de la Universidad Pompeu Fabra, del cual fue director hasta el año 2005. Sus principales líneas de investigación y docencia son la hacienda pública, en sus diferentes vertientes analíticas y aplicadas o de gestión. Entre otras responsabilidades, es codirector del Máster de Gestión Pública (UPF-UAB-Esade) y miembro de diferentes consejos asesores y editoriales de revistas de ámbito nacional e internacional. Desde el año 2005 es Consejero del Banco de España.

Mancebón Torrubia, Mª Jesús. Licenciada en Ciencias Empresariales y Doctora en Economía por la Universidad de Zaragoza. En la actualidad es Profesora Titular de Eco-

Sumario

nomía Aplicada en la Facultad de Ciencias Económicas y Empresariales de la Universidad de Zaragoza. Sus principales líneas investigadoras se concretan en el estudio de la eficiencia productiva en diversos ámbitos, principalmente en el sector educativo pero también en el área de los servicios policiales y del sector industrial español. Ha publicado en diversas revistas tales como *Journal of the Operational Research Society, Education Economics, Applied Economics, Hacienda Pública Española, Información Comercial Española* y *Papeles de Economía Española*.

MARTÍNEZ PÉREZ, Jorge Eduardo. Su tesis doctoral versa sobre la valoración monetaria de las reducciones del riesgo de muerte, y su investigación actual se centra en aspectos relativos a la evaluación económica en economía de la salud. Ha impartido docencia en la Universidad Politécnica de Cartagena, y actualmente es miembro del departamento de Economía Aplicada de la Universidad de Murcia.

MURCIANO ROSADO, Juan. Investigador del Instituto de Desarrollo Regional, Fundación Universitaria. Licenciado en Economía por la Universidad de Sevilla. Experto universitario en evaluación de programas y políticas públicas por la Universidad Complutense. Coordinador del programa «Horizonte 2007» para la promoción de la cultura de evaluación en el Pacto Territorial por el Empleo de la Bahía de Cádiz. Miembro de la Sociedad Española de Evaluación. En la actualidad es investigador del Instituto de Desarrollo Regional adscrito a la línea de Evaluación de Políticas Públicas desde octubre 2003. Participa en la asistencia técnica a la Consejería de Empleo de la Junta de Andalucía para el seguimiento y evaluación de la Iniciativa Comunitaria Equal en Andalucía.

ONRUBIA FERNÁNDEZ, Jorge. Economista y Doctor en Economía por la Universidad Complutense de Madrid. Profesor Titular de Hacienda Pública en la Facultad de Ciencias Económicas y Empresariales de esa Universidad, en la que enseña Economía Pública desde 1988. Ha sido investigador en plantilla del Instituto de Estudios Fiscales desde 1991 hasta 2004. Sus áreas de investigación son el análisis económico de la gestión pública, la economía de la imposición y el análisis insitucional de la política fiscal, acumulando un número considerable de artículos en revistas nacionales y extranjeras. Ha participado en varios proyectos de reforma presupuestaria y organizativa en la Administración del Estado.

ORTÚN RUBIO, Vicente. Profesor del Departamento de Economía y Empresa y Director del Centro de Investigación en Economía y Salud (CRES) de la Universidad Pompeu Fabra, Barcelona. Director de la colección de libros sobre 'Economía de la Salud y Gestión Sanitaria': 18 títulos publicados por Springer-Verlag, Masson y Elsevier. Director del Master en Economía de la Salud y Gestión Sanitaria de la UPF con ediciones en Barcelona y Madrid. Profesor del Health Leadership Forum, conjunto con la Universidad de Berkeley. Es Máster of Science por Purdue University, Licenciado y MBA por ESADE, Licenciado y Doctor en Ciencias Económicas por la Universidad de Barcelona y ha reali-

Sumario

zado estudios de doctorado en Salud Pública en Johns Hopkins University. Experiencia en gestión sanitaria nacional (Ministerio Sanidad y Consumo, Generalidad provisional de Cataluña, Hospital de San Pablo en los primeros años setenta...) e internacional. Ex-presidente de la Sociedad Española de Salud Pública y Administración Sanitaria (SESPAS). Ex-presidente de la Asociación de Economía de la Salud (AES). Ex-secretario de la European Public Health Association (EUPHA). Fundador de la Sociedad Española de Epidemiología, la Internacional Health Policy Association, la Fundación Avedis Donabedian, SESPAS, AES, y EUPHA. Autor de 300 artículos y libros, nacionales e internacionales.

Osuna Llaneza, José Luis. Director del Departamento de Economía Aplicada II. Universidad de Sevilla. Doctor en Ciencias Económicas por la Universidad de Sevilla. Director de programas de la línea de Evaluación de Políticas Públicas del Instituto de Desarrollo Regional. Es Catedrático E. U. de la Universidad de Sevilla y Director del Departamento de Economía Aplicada II de la Facultad de Ciencias Económicas y Empresariales de la Universidad de Sevilla. Ha sido Stagiare Administratif de la DG XIV, Comisión Europea, y actualmente es miembro del Consejo Asesor para la Ciencia y la Tecnología en Andalucía, del Consejo Económico y Social de Andalucía y del Comité de Expertos para la creación de la Agencia Estatal de Evaluación de la Calidad de los Servicios y de las Políticas Públicas. Desde 2005 es el Vicepresidente de la Sociedad Española de Evaluación. Experto en Economía Regional, ha publicado libros y artículos sobre distintos aspectos de la estructura productiva andaluza y sus implicaciones en las políticas de la Unión Europea y ha dirigido numerosos trabajos de evaluación.

Pinto Prades, José Luis. Licenciado en Economía por la Universidad de Valencia y postgrado en Economía de la Salud por la Univesidad de York (Reino Unido). Se doctoró en Economía por la Universidad de Murcia. Actualmente trabaja en la Universidad de Pompeu Fabra (barcelona). Sus primeras investigaciones están relacionadas con cuestiones teóricas y prácticas acerca de la Evaluación Económica de las tecnologías relacionadas con el cuidado de la salud. Ha sido asesor del Ministerio de Sanidad español en evaluaciones económicas de tecnologías de la salud. También ha colaborado con la OCDE en grupos de trabajo acerca de tecnologías emergentes relativas a la salud (2001-2003) siendo uno de los representantes expertos en el Ministerio de Sanidad español. Ha publicado en revistas internacionales como *Health Economics, Journal of Health Economics, Medical Decision Making,* y *Management Science and the Economic Journal.*

Planas Miret, Ivan. Es profesor asociado del Departamento de Economía y Empresa de la Universidad Pompeu Fabra. Investigador del Centro de Investigación en Economía y Salud (CRES). Licenciado en Ciencias Económicas y Empresariales por la Universidad Pompeu Fabra en 1996. Máster en Economía Pública por la Universidad de York (R.U.) en 2000. Sus principales líneas de investigación y docencia son la economía de la salud, la economía pública y la gestión pública.

Sumario

SUBIRATS HUMET, Joan. Doctor en Ciencia Económicas, catedrático de Ciencia Política y Director del Instituto de Gobierno y Políticas Públicas en la Universidad Autónoma de Barcelona. En el año 2003 ocupó la cátedra Principe de Asturias en la Universidad de Georgetown. Se ha especializado en temas de gobernanza, gestión pública y en el análisis de políticas públicas y exclusión social, habiendo publicado diversos libros relacionados con ello. Ha estudiado e investigado en problemas de innovación democrática, sociedad civil y gobierno multinivel. Colaborador habitual en diversos medios de comunicación.

VELEZ MENDEZ, Carmen. Investigadora del Instituto de Desarrollo Regional, Fundación Universitaria. Licenciada en Ciencias Económicas y Empresariales. Coordinadora de la línea de Evaluación de Políticas Públicas del Instituto de Desarrollo Regional, ha realizado trabajos de evaluación tanto para la Unión Europea como para la Administración Central española, el gobierno regional de Andalucía y distintas administraciones locales (ayuntamientos, mancomunidades y diputaciones). Ha participado en la redacción de la *Guía para la Evaluación de Políticas Públicas* así como de la *Guía de Planificación y Evaluación para Agentes de Desarrollo Local* editada por Sevilla Siglo XXI. Actualmente es vocal de la Sociedad Española de Evaluación y fue miembro fundacional de la misma. Es profesora asociada en el Departamento de Organización Industrial y Gestión de Empresas de la Escuela Superior de Ingenieros Industriales de Sevilla. Fue investigadora visitante en el Center for Regional Economic and Social Research de la Sheffield Hallam University. Actualmente realiza su tesis doctoral sobre evaluación de políticas públicas.

Sumario

Sumario

EKONOMIAZ

ÚLTIMOS NÚMEROS PUBLICADOS

30. Economía y Empresas
31/32. El Reto del Empleo
33. Economía de la Participación
34. El Reparto de Trabajo
35. El Debate Económico Actual
36. Comercio Exterior: Tendencias Recientes
37. Competencia y Regulación de Mercados
38. El Dilema Fiscal
39. Medir en Economía. Lo que sabemos que no sabemos
40. Economía e Igualdad
41. Economía y Territorio
42. El Futuro del Estado de Bienestar
43. Capital Humano y Transiciones Laborales de los Jóvenes
44. Competitividad y empleo en la economía vasca
45. La Economía del conocimiento
46. Panorama de las industrias de red
47. La empresa innovadora
48. Globalización y sistema financiero: los desafíos ante el nuevo milenio
49. Medio ambiente y recursos naturales: nuevas aportaciones
50. El gobierno de la empresa
51. La propiedad intelectual en la sociedad de la información
52. Lecciones de la deflación: Estados Unidos frente a Japón
53. La política de *clusters* en el País Vasco
54. El siglo xx en la historia económica del País Vasco: de la gran empresa a las pymes
55. Inversión extranjera directa y procesos de deslocalización
56. Ciencia, tecnología, innovación y sociedad
57. Valoración de activos ambientales: la catástrofe del *Prestige*
58. Ciudades Región Globales. Espacios Creativos y Nueva Gobernanza
59. Capital Social. Innovación Organizativa y Desarrollo Económico

Sumario

Sumario

BOLETÍN DE SUSCRIPCIÓN

Nombre o Razón Social. . . . . .
Domicilio . . . . . .
Localidad . . . . . D.P. . . . . .
Provincia o País. . . . . .
DNI/NIF . . . . . Tfno. . . . . .

TARIFAS: Suscripción anual (3 números)

- Particulares: 18 €.
- Instituciones y empresas: 30 €.
- Números sueltos: 12 €.
- Estudiantes 20% de descuento

PAGO ☐ Talón bancario a favor de «Tesorería General del Gobierno Vasco»
☐ Transferencia bancaria

Firma o Sello

EKONOMIAZ

Ogasun eta Herri Administrazio Saila. Ekonomia eta Plangintza Zuzendaritza
Dpto. de Hacienda y Administración Pública. Dirección de Economía y Planificación
Eusko Jaurlaritza/Gobierno Vasco. Donostia-San Sebastián, 1 – 01010 Vitoria-Gasteiz
Tfno.: 945 019 038 – Fax: 945 019 062
E-mail: economia@ej-gv.es
http://www1.euskadi.net/ekonomiaz/indice_c.htm

Sumario

HARPIDETZA-ORRIA

Interesatuaren edo Sozietatearen Izena . . . . . .
Helbidea. . . . . .
Herria . . . . . D.B. . . . . .
Probintzia edo Herrialdea . . . . . .
NAN/IFZ . . . . . Tlf. . . . . .

TARIFAK: Urteko Harpidetza (3 ale)

- Bereziak: 18 €.
- Erakunde eta enpresak: 30 €.
- Ale solteak: 12 €.
- Ikasleentzat %20ko beherapena

ORDAINKETA ☐ Banku-taloia «Eusko Jaurlaritzako Altxortegi Orokorra»ren alde
☐ Banku-transferentzi

Sinadura edo zigilua

EKONOMIAZ

Ogasun eta Herri Administrazio Saila. Ekonomia eta Plangintza Zuzendaritza
Dpto. de Hacienda y Administración Pública. Dirección de Economía y Planificación
Eusko Jaurlaritza/Gobierno Vasco. Donostia-San Sebastián, 1 – 01010 Vitoria-Gasteiz
Tfno.: 945 019 038 – Fax: 945 019 062
E-mail: economia@ej-gv.es
http://www1.euskadi.net/ekonomiaz/indice_c.htm

Sumario

Fundamentar cambios en políticas existentes o introducir nuevas estrategias exige el mejor conocimiento, a ser posible de decisiones basadas en la evidencia disponible. Dicho recurso implica usar instrumentos que coadyuven a dichos cambios. En ningún caso se trata de sustituir la decisión política responsable en la materia, lo que se intenta es informar mejor acerca de sus consecuencias. De ahí que la gestión y evaluación de las políticas públicas no se puede permitir viajar sin alforjas. Por ello la utilidad, la calidad y la idoneidad de los instrumentos de evaluación de políticas públicas en el contexto actual, no pueden limitarse además a ser meros controladores de la legalidad o a detectar abusos de poder. Deben ir más allá y convertirse en indicadores reales de impacto además de reflejar la calidad de las interrelaciones entre los organismos públicos y el exterior.

EUSKO JAURLARITZA GOBIERNO VASCO

Euskal Jaurlaritzaren Argitalpen Zerbitzu Nagusia
Servicio Central de Publicaciones del Gobierno Vasco

P.V.P. (2 vol.): 12 €