

# ACERCA DE LA MEDICIÓN Y ANÁLISIS DE LA EFICIENCIA EN LAS ORGANIZACIONES SANITARIAS PÚBLICAS\*

**Guillem López i Casasnovas**

*Universitat Pompeu Fabra*

En este artículo se abordan algunos desarrollos recientes en el análisis de la eficiencia en los servicios sanitarios y sus aplicaciones, especialmente en el campo de la presupuestación. Se propone un ejercicio de presupuestación alternativo basado en el análisis de la eficiencia y en la presupuestación prospectiva; para analizar, a continuación, las aportaciones que desde la Economía de la Salud se han hecho al estudio de la eficiencia en el suministro de los servicios sanitarios y de los problemas y métodos que se plantean en dicho ámbito. Finaliza el trabajo vinculando los aspectos relativos al análisis de la eficiencia con los problemas de presupuestación. Se señalan, asimismo, algunas de las limitaciones del análisis realizado en su aplicación a la gestión.

*Palabras clave:* eficiencia, servicios sanitarios, presupuestación.

## 1. INTRODUCCIÓN

En este trabajo nos proponemos comentar algunos desarrollos recientes en el análisis de la eficiencia en los servicios sanitarios y las virtualidades de sus aplicaciones, en particular en el campo de la presupuestación de los servicios.

El trabajo se divide en dos partes. En la primera se identifican los problemas de financiación sanitaria al nivel de presupuestación de organizaciones y de servicios. En la segunda, se revisa la contribución de la literatura en Economía de la Salud al análisis del grado de eficiencia con la que se suministran los servicios sanitarios, y se engarzan los resultados de dicho análisis con la problemática existente en los temas de financiación, señalando los puntos fuertes y débiles de dicha conexión.

---

(\*) Este trabajo se ha beneficiado de la financiación de la Dirección General de Investigación Científica y Técnica (DGICYT) del Ministerio de Educación.

## 2. LA PRESUPUESTACIÓN DE LAS ORGANIZACIONES SANITARIAS COMO FINANCIACIÓN POR LA COMPRA DE SERVICIOS

Uno de los problemas que más consecuencias tiene para una buena gestión de los servicios públicos es el asociado a la falta de motivación para la eficiencia que generan los procesos de presupuestación. Dichos procesos contienen factores de universalismo (todos los agentes esperan obtener algún beneficio en el proceso), reciprocidad (intercambio de apoyos) e incrementalismo (se discute el ritmo de crecimiento y no la base de partida). Como resultado, en muchas ocasiones, el presupuesto es poco más que un simple reembolso, por parte de la autoridad financiadora, de los costes de suministro del servicio.

En efecto, y pese a lo que se pretende a menudo aparentar, de hecho resulta difícil que el financiador se aleje de la inercia que suponen los gastos ya registrados en la realidad. Este es en particular el caso cuando los gastos derivan de una estructura jerárquica supuestamente controlada por el propio financiador, y son resultado de ineficiencias que el responsable no puede aducir que ignora, sino que en todo caso habrá de postular. Así, la restricción presupuestaria en la vida de las organizaciones públicas resulta laxa y más ligada a criterios de coste histórico que de eficiencia.

Por lo demás, en un contexto de financiador único y de múltiples unidades gestoras de servicios, atender a dichos costes de un modo uniforme, sin mayor corrección, puede generar agravios comparativos, y forzar, en consecuencia, un ajuste financiero indiscriminado, independientemente de la mayor o menor eficiencia que pueda contener el gasto realizado por el ente gestor.

### *Una óptica alternativa*

Para un ejercicio de presupuestación alternativo, el análisis de la información disponible sobre la eficiencia relativa en la gestión de los distintos servicios resulta fundamental. Una aproximación simple a lo anterior consiste en sistematizar las diferencias a través de los denominados procesos de evaluación por comparación. Se trata de eliminar los elementos comunes para centrarse en los idiosincráticos propios de cada institución. Dicha estrategia cobra un mayor interés cuando efectivamente se comparan comparables: por ejemplo, en aquellos casos en que distintas unidades de provisión pública han de suministrar en lugares diferentes de la geografía, en principio, estándares de servicio similares. Si la descentralización orgánica se acompaña verdaderamente de autonomía en la gestión, la virtualidad de las políticas de evaluación relativa pueden contribuir al conocimiento de la eficiencia del gasto a la hora de su presupuestación.

Para ello, a diferencia de lo que acontece con la disponibilidad de la información privada, que dado el valor de su contenido sobre la actuación de los agentes deviene rival, la información en la gestión pública adquiere, en el contexto anterior, características de bien público: sus externalidades son positivas y añaden valor al bienestar colectivo. A su vez, dicho conocimiento supone una mayor transparencia en la gestión, ya que aminora la asime-

tría de información en general existente entre financiador y entes financiados<sup>1</sup>.

### *2.1. Competir por fondos en contraprestación de actividad*

La importancia de evaluar el presupuesto sobre bases fundadas en costes y en la actividad requerida de las instituciones suministradoras de servicios por parte del financiador es crucial en aquellos casos en los que no se dispone de información relativa a las preferencias y/o disponibilidades de pago de los ciudadanos.

En un contexto como el anterior, la política de «bench-marking» en la fijación de estándares de gasto o una política de «torneos» en la consignación de fondos según criterios de eficiencia relativa puede señalar parte del camino a recorrer en la gestión sanitaria pública.

A dichos efectos, una primera posibilidad consiste en la construcción de un estándar de utilización normada de servicios (a la vista de la tipología de usuarios reales) o de disponibilidad de recursos. Este último podría entenderse como capacidad de respuesta (accesibilidad) de un potencial usuario al servicio, a partir de aquellas variables externas que, en el conjunto de la muestra, influyen en el nivel de provisión de un servicio. Este tipo de estándares puede ser de particular interés para aquellas actuaciones lideradas por la demanda, pero para las que no se desee reconocer la totalidad del consumo observado como necesario con criterios de eficiencia social. Así, un indicador de frecuentación (utilización) media de un servicio, ponderado a un nivel de coste medio o estándar, puede señalar fácilmente un criterio de gasto, y presupuestación consiguiente para dicho servicio, evitando el reembolso automático de cualesquiera hayan sido los recursos consumidos.

El cuadro 1 remite al lector a una tabla resumen de modalidades de presupuestación, manteniendo la idea básica de las políticas comentadas; esto es, de financiar el gasto correspondiente a la actividad legítimamente realizada, y/o financiarla asumiendo tan sólo aquellos niveles de coste estimados eficientes, evitando un resarcimiento indiscriminado de cualesquiera sean los niveles de coste observados.

Nótese que en los casos en que la actividad no se estandariza, pero sí los costes de eficiencia, el seguimiento de la calidad ha de poder considerarse realizado a partir de la observación de cómo eligen los ciudadanos los distintos centros productores de servicios.

En esta aproximación, la tipología de casos atendida, vista la diversidad de intervenciones, puede ajustarse con una consignación financiera, especí-

---

(1) De hecho, algunas instituciones privadas intentan un resultado pareto superior actuando como fedatarios de información privada. Su actuación suele consistir en la agregación, a distintos niveles, de la información remitida por los particulares, a su vez destinatarios, de la información. Sin permitir ser identificados particularmente, ni identificar a terceros, esta elaboración de información sirve, en principio, para un posicionamiento relativo por ámbitos, sectores o ramas de la actividad económica, a las empresas. El Banco de España, por ejemplo, actúa como «broker» de información para las empresas privadas respecto de los ratios económicos más básicos.

**Cuadro 1**  
**GASTO PRESUPUESTARIO (G), FINANCIACIÓN PÚBLICA (F) Y**  
**COMPETENCIA POR LOS FONDOS PRESUPUESTARIOS DISPONIBLES**

---

GASTO  $i = P_i \times Q_i$

---

Con referente  $Q_i$  (cantidad observada pero a coste regulado)

P (CMe)

P (CMg)

P (CMe\*)

---

Con referente  $P_i$  (a coste observado, pero utilización «normada»  
o abierta —resultado de la elección)

Q (Qe)

Q (Q\*)

---

Vinculación entre gasto y financiación

$(P \times Q)_i : F_i = G^* (P^*, Q^*)$

$(P \times Q)_i : F_i = G (P (CMe), Q_i)$

$(P \times Q)_i : F_i = G^* (P^*, Q_i)$

---

Notas: P indica «precio» o coste unitario; Q refiere el nivel de actividad; (\*)=Predicción;  
CMe= coste medio; CMg=coste marginal; i=destinatario de la financiación;  
e=actividad resultado de una elección.

---

fica o incluida en el monto global, que atienda a grupos funcionales homogéneos (grupos idiosincráticos similares), imputando a éstos costes presupuestarios idénticos.

Otra posibilidad consiste en centrar el análisis en el refinamiento del estudio de costes necesarios («legítimos»), a través de la medición de la eficiencia de la actividad pública, aceptando, sin embargo, que el nivel de actividad observado es el necesario; por ejemplo, al no considerarse éste manipulable ni por parte del financiador ni del ente gestor.

Finalmente, para aquellas actividades preventivas, en las que el mejor resultado es evitar la necesidad de utilización de un servicio, la estimación de los factores condicionantes de la necesidad de la actividad y el diferencial consiguiente entre dicha previsión y la realmente observada, puede ofrecer una medida aproximativa al output de un servicio.

## 2.2. De la evaluación de la eficiencia a la financiación

El paso del registro de la evaluación de lo realizado a la presupuestación prospectiva parece conveniente que se explicita como si de un proceso de retro-alimentación se tratase.

Sea la siguiente función de costes:

$$C = C(Q, R(z), w)$$

en la que Q supone actividad o vector de outputs, w precio de los inputs y R un conjunto de factores exógenos relacionados, por ejemplo al riesgo z de la estructura de oferta servida.

Si la autoridad sanitaria pública determina unos precios de los servicios  $p^*$ , en favor de una compra eficiente de servicios, y la empresa suministradora intenta maximizar sus ingresos totales:

$$TR = p^* \cdot Q$$

restringida a la función de costes anterior, las condiciones de primer orden conducen a precio igual a coste marginal

$$p^* = C_1(Q)$$

Sin embargo, este es el caso sólo cuando  $Q$  es independiente de  $R$ .

Siendo la función de costes medios

$$C(q) = C(R(z), v)$$

y  $R(z)$  un vector de variables explicativas exógenas y  $v$  el término de error.

Si los precios regulados  $p^*$  resultan inferiores a los costes medios  $AC$  ello se podrá deber a (i) que las variables de ajuste  $R$  generan diferencias en el coste,<sup>2</sup> o (ii) una menor eficiencia  $v^*$ . La autoridad pública desea posiblemente compensar por el primer factor, pero posiblemente no reconocer en su totalidad el segundo, en la medida que éste haya sido resultado de una decisión discrecional o arbitraria por parte del suministrador<sup>3</sup>.

En este contexto, y siguiendo a Schleifer (1985) si todas las unidades financiadas son idénticas:

$$p^* = AC$$

en otro caso (unidades heterogéneas), la propuesta formulada por Pope (1990) fijaría

$$AC_i^* = b(AC_i) + (1-b) AC$$

debiendo  $b$  ser mayor cuanto más alta fuese la proporción de la varianza total debida a costes justificables, y mayor fuese la correlación de dichos factores con los costes observados.

Y ponderando la aplicación de la fórmula anterior en el tiempo (Ellis, McGuire, 1990).

$$AC_i^* = b^t AC_i + (1-b^t) AC$$

$$0 < b < 1 \quad t = 1, \dots, n.$$

Remitiendo la formulación anterior al contexto sanitario, toda tarifa situada por debajo de los costes medios «justificables» o esperados por parte del destinatario de la financiación, ya sea éste una aseguradora o un centro sanitario, puede agravar una reacción no inequívocamente relacionada con la in-

(2) Por ejemplo, de acuerdo con cierta selección de riesgos o favoreciéndose de los entornos de actividad más favorable en costes (menores costes de acceso, posibilidad de economías de escala o alcance, etc.).

(3) De ahí la importancia de descomponer el término residual en la estimación de los costes entre el componente aleatorio y la ineficiencia.

centivación a la eficiencia; por ejemplo, puede motivar una mayor selección favorable de riesgos por la que se ingresen tan sólo pacientes de baja complejidad o se aseguren individuos de riesgo previsible bajo. Como consecuencia, el objetivo perseguido de presionar vía tarifas sobre la ineficiencia no resulta garantizado.

Utilizar, como veíamos, una tarifa prospectiva definida sobre la base de los costes medios del conjunto de destinatarios de la financiación sólo podrá ser óptimo si se trata de unidades idénticas en todos los aspectos relevantes que influyen en los costes «legítimos» o que se consideran «justificables» para la realización de una determinada actividad. Si ello no fuese el caso, la financiación óptima habría de recoger, como tarifa, un valor medio ponderado de los costes unitarios propios y del conjunto de unidades con los que se compare o compita. No resulta sorprendente, por lo demás, que el peso de los costes propios haya de ser mayor cuanto mayor sea la proporción de la varianza total atribuible a factores justificables y mayor la correlación de éstos con los costes observados.

La definición de costes «justificables» a efectos de la determinación de contratos óptimos se puede basar en los ajustes explícitos de riesgos predecibles fuera del alcance de la actividad de la unidad financiada en su impacto sobre la utilización de los factores, o a través de ajustes en los costes unitarios, que recojan factores exógenos a la gestión del destinatario de la financiación. En otras palabras, ello se puede conseguir definiendo una noción de «costes predichos», o costes en los que debiera de haber incurrido el prestador en caso de haberse ajustado a determinadas pautas de comportamiento.

El concepto de costes «justificables» puede derivarse, por ejemplo, de la estimación de funciones frontera —de producción o de coste, paramétricas o no, determinísticas o estocásticas<sup>4</sup>—. Asimismo, una aplicación de la teoría de la «*yardstick competition*» de Shleifer antes referida, en supuestos de no observabilidad del esfuerzo, puede derivarse de la literatura de «torneos» con ordenaciones de coste y actividad que generen una discriminación entre unidades eficientes y no eficientes, de acuerdo con alguna regla preestablecida, y suponiendo que se produce en cada unidad la compensación interna de los elementos aleatorios que puedan estar presentes.

Si la base de competencia es muestral, de todos los concurrentes, el participante puede influir en el resultado, pero es el conjunto de la muestra quien decide.

Tal como han puesto de manifiesto recientes trabajos sobre modelos mixtos de financiación (Newhouse, 1994)<sup>5</sup>, el peso que óptimamente deben de tener los costes propios de la unidad financiada en la fórmula de presupuestación depende, positivamente, del coeficiente observado entre la variación en el tipo de actividad que realiza el centro y la variación en los costes estan-

---

(4) Un resumen de dicho tipo de aproximación se encuentra en el trabajo de Schmidt y Lovell (1979).

(5) Véase sobre el tema la revisión que sobre la literatura anterior realiza Newhouse J. P. (1994a).

darizados (por ejemplo, en los valores medios de las unidades) y, a la vez, de la correlación que muestre la complejidad y tipología de actividad y los costes estandarizados. Es decir, cuanto más elevada sea la desviación estándar en la complejidad respecto a la correspondiente a los costes medios, mayor es la importancia que en la fórmula de reembolso del centro debe de tener sus costes específicos. Esta es la propuesta hoy de las formas mixtas de financiación (*blended*). La correlación entre costes y caracterización de la actividad actúa, en este sentido, como una simple ponderación del coeficiente anterior.

A la vista de lo anterior, sobre el ordenamiento resultante de los valores teóricos y los actuales, y aún aceptando que existan deficiencias en la información siendo, sin embargo, éstas comunes a la muestra y no específicas a cada uno de los participantes, la política de torneos podría utilizarse para la promoción de entornos competitivos en gestión pública<sup>6</sup>.

Para objetivar los premios y castigos vinculados a los resultados de la competencia —a quién se incentiva positivamente y a quién se penaliza— y evitar el efecto trinquete, o rebote, que supondría incorporar secuencialmente las mejores prácticas observadas de un modo discrecional por parte del financiador, una posibilidad consiste en definir la cuantía de los agentes afectados en cada uno de los supuestos anteriores (promoción-relegación) de un modo porcentual: esto es, prefijando un porcentaje de unidades de la muestra a efectos de compensación positiva o negativa. Ello se puede concretar, en la práctica, si asumimos una distribución normal para el indicador de eficiencia elegido, a través de su determinación en términos del valor observado de la media (dado que todos los participantes influyen en ella) más un número determinado de desviaciones típicas. Una aproximación de este tipo objetiva el sistema de promociones y sanciones, alejándolo de la discrecionalidad del financiador.

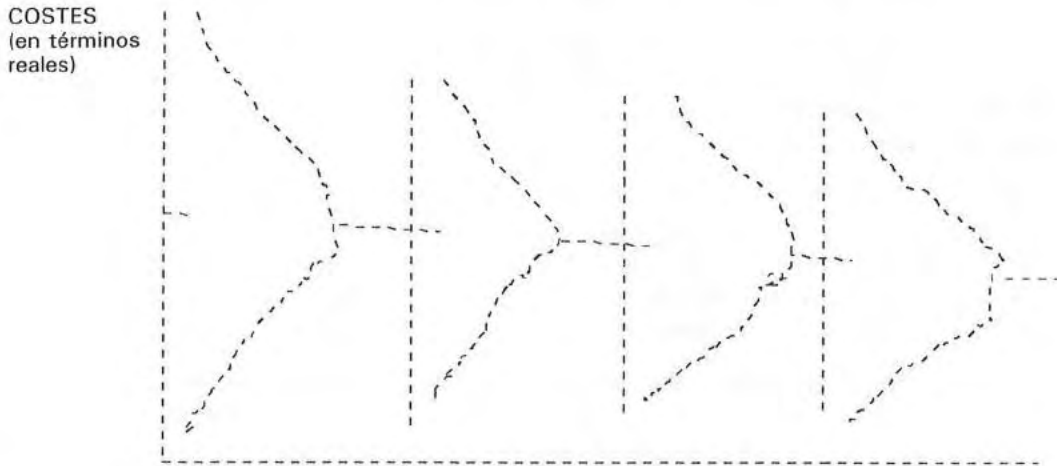
La operativa en el tiempo de un sistema como el comentado se refleja en el gráfico adjunto. Como puede observarse, incentivar que se dirijan hacia la media las desviaciones negativas en actividad o positivas en costes medios, conlleva, de acuerdo con los supuestos comentados, a que, en términos reales, disminuya la ineficiencia media y se concentre la distribución en torno a dicho valor.

#### *Períodos temporales en los que opera el sistema*

En la aproximación hasta ahora comentada, la definición de estándares de eficiencia, ya sea por el de la valoración de actividad o de costes, resulta crucial. Los desarrollos recientes en investigación de servicios sanitarios redirigen hacia la estimación de la eficiencia sanitaria un importante núcleo de trabajos [desde las aportaciones iniciales de M. Feldstein (1975), entre otros]. En este contexto, resulta natural interrogarse sobre si los resultados de éstos pueden contribuir a la construcción de la estrategia comentada en esta sección, a fin de contemplar la financiación a los centros sanitarios como el resultado de una compra de servicios regida con criterios de eficiencia. A dicho tema se dedica el segundo gran apartado del presente texto.

(6) Véase en este extremo Macho I. y Pérez-Castillo D. (1994): *Introducción a la Economía de la Información*. Ed. Ariel.

**Gráfico 1**  
**ILUSTRACIONES DE LA OPERATIVA PROPUESTA**



### 3. LA MEDICIÓN DE LA EFICIENCIA DE LOS SERVICIOS SANITARIOS

Desde la aparición de los primeros estudios sobre el tema de la medición de la eficiencia hospitalaria (López-Casasnovas, 1984) se ha producido en España una verdadera avalancha de trabajos que se centran en distintos aspectos de la medición de la eficiencia de los servicios sanitarios<sup>7</sup>. Nuestro trabajo inicial era econométricamente sencillo y limitado en sus aspiraciones. Dicho en otras palabras, el objetivo del análisis no era «pagar» en función de los resultados de la estimación de la eficiencia, ni de incentivar la actividad con sistemas de productividad basada en productos intermedios. Se trataba tan sólo de contar con elementos diversos que trataran de sacar de la «indigencia» evaluativa al financiador frente a los requerimientos de presupuestación por parte de los entes gestores del gasto. En este sentido, cada centro, si bien es distinto en algunos factores para lo que ajustamos rasgos estructurales y elaboramos índices de complejidad y especialización, contiene rasgos comunes que permiten su comparación. Las diferencias, una vez realizado lo anterior se trasladan, en cuanto a la carga de la prueba de su justificación, al ente financiado.

A medida que más y mejores estadísticas han aparecido, sobre todo en literaturas relacionadas, su importación ha permitido sofisticar más y más las estimaciones y aumentar sus supuestas potencialidades prescriptivas. Sin embargo, ello no siempre se ha visto acompañado de una validación previa de la información, y sí muchas veces de la simplificación de los postulados del análisis sin acotar con ello poder prescriptivo para sus resultados. Además, el interés, en algunas ocasiones, por mejorar las estimaciones ha generado

(7) Bohigas, Puig, González-Valcárcel, Barber, López Nicolás, Quintana, Rué, Prior, Solá, Sáez, Urbanos, entre otros, se han ocupado de aplicar distintas técnicas a diversos ámbitos del suministro de servicios sanitarios (hospitales, centros de salud, unidades de cuidados intensivos, procedimientos quirúrgicos, etc.).



especificaciones más o menos «ad hoc»: desde suponer que se minimizaban los costes para poder utilizar formas funcionales flexibles, a utilizar como precios los costes unitarios de producción para ajustar mejor el ajuste predictivo.

Todo ello crea un cierto vértigo al estudioso de dichas materias que hace conveniente una especie de parada y fonda para comentar algunos de los problemas asociados a dicha aproximación, y determinadas líneas de mejora, sobre todo si se quiere saltar del ejercicio intelectual al plano del «policy making».

### 3.1. Algunos problemas

#### A. La cuestión de los datos

Un primer problema, que pasaremos de «puntillas», reside en la aplicación estadística no a veces a muestras poblacionales sino a la población en su conjunto. Son muchos los trabajos, en este sentido, que utilizan datos, por ejemplo, de todos los hospitales, sin explicitar dicho extremo a efectos de estimación.

Por lo demás, si se define una muestra, para objetivos de evaluación por comparación, hace falta comprobar, sin sesgos, que se comparan comparables. Muchos estudios sitúan en la misma estimación hospitales de unas pocas camas, digamos cincuenta, con grandes hospitales, digamos mil doscientas. De modo similar en lo que se refiere a encuadrar en la misma muestra centros de agudos y larga estancia, próximos a ciudades sanitarias, con hospitales rurales aislados, maternidades separadas o integradas en el complejo sanitario, con hospital de día o no, etc. En la medida en que dichos elementos diferenciales condicionen la eficiencia en gestión, y no sean directamente imputables al gestor (son resultados de la planificación pasada o de opciones de política sanitaria), se requiere un ajuste previo. Ello se puede efectuar ya sea segmentando la muestra en submuestras o introduciendo variables específicas en la estimación que recojan aquellos factores. Hacer una u otra cosa no resulta neutro ya que, en el segundo caso, se mantiene una idéntica especificación funcional y se asume idéntica estabilidad estructural, limitando el impacto de la diferente naturaleza del centro a la variable explicativa en cuestión.

En lo relativo a centros de agudos, vistos según tamaño, algunos estudios intercambian camas institucionales o autorizadas con camas ocupadas como si de lo mismo se tratase, e incluso confundiendo entre ellas las neonatales. Por lo que atañe a la actividad, la progresiva mejora de los datos (gracias a la implantación reciente del Conjunto Mínimo Básico de Datos - CMBD) hace que algunos trabajos, que combinan información en datos de panel (de corte transversal y serie temporal) fijen de manera estática la variable de casuística, es decir, al no estar disponible ésta para cada año, sobre la misma base de información, se asume su constancia a un nivel de agregación temporal mayor.

#### B. Las variables incluidas y excluidas en la estimación y la interpretación consecuente de los resultados

El modo en el que se concreta la ecuación a estimar, de nuevo va a tener efectos en la interpretación que se podrá después dar a los resultados: por

ejemplo, ante el análisis de las variaciones de la actividad hospitalaria (básicamente, internamiento de pacientes) en los costes de los hospitales. Por ejemplo, en cuanto a valorar el impacto que supondría: a) una variación en la estancia media hospitalaria de una manera aislada, o b) cuando aquélla se acompañase de distintos niveles de utilización de la infraestructura hospitalaria.

Pongamos el caso de la siguiente estimación, a modo de ejemplo<sup>8</sup>:

$$\begin{aligned}
 (1) \quad \text{Coste por paciente} &= 53,32 - 1,165 \text{ FL} + 4,38 \text{ EM} - 0,721 \text{ DVMIR} \\
 &\quad (\text{en miles de ptas.}) \quad (1,62) \quad (-2,99) \quad (2,53) \quad (-0,192) \\
 &\quad - 10,02 \text{ DVCE} - 0,458 \text{ MIECST} + 1,532 \text{ CGECST} \\
 &\quad \quad \quad (-1,753) \quad \quad (0,20) \quad \quad (-0,27) \\
 &\quad R^2=0,64 \quad \bar{R}^2=0,56 \quad F(9,40)=7,868
 \end{aligned}$$

Donde FL=N.º enfermos ingresados al año por cama disponible. EM=Estancia media. DVMIR=Variable que recoge el efecto de la docencia (variable dicotómica igual a 1 en caso de hospital con MIR, e igual a cero en otro caso). DVCE=Variable que recoge el efecto de las consultas externas (variable dicotómica para la provisión de consultas externas, siendo igual a 1 si el ratio de pacientes tratados en relación al número de pacientes ingresados en el hospital muestra valores superiores a la media del grupo de hospitales considerados, siendo 0 en otro caso). MIECST=Proporción sobre el total de casos tratados en la especialidad de medicina interna.

CGEST=Idem, para Cirugía General y especialidades quirúrgicas. OBGIST=Idem, para el servicio de Obstetricia y Ginecología. PECST=Idem, para Pediatría. CUICTST=Idem, para casos tratados en Unidades de Cuidados Intensivos. (Los valores se expresan en miles de ptas.).

La ecuación (1) permite, en efecto, cuantificar el efecto marginal combinado de variaciones en distintas variables de utilización. Para ello conviene recordar la interpretación convencional dada a los coeficientes estimados de forma aislada, y el hecho de que las propias variables de utilización hospitalaria más comúnmente utilizadas se hallan ligadas entre sí de acuerdo con la relación:

$$\text{EM} = \frac{\text{Coef. de utilización}}{\text{Admisiones por cama}} \times 365$$

debido a que el coeficiente de utilización se define como el total de estancias dividido por el número de camas disponibles multiplicado por 365 días (o estancias posibles).

De este modo, la inclusión de dichas variables en la estimación permite la interpretación de los coeficientes referidos a cada una de las variables, una

(8) Notas: La estimación ha sido efectuada por Mínimos Cuadrados, ponderados por el tamaño del hospital (número de camas disponibles), debido a la no superación en una primera aproximación del test de heteroscedasticidad de los residuos.

Los valores entre paréntesis reflejan la t de Student, considerándose «aceptable» en la significatividad del coeficiente estimado si registra valores superiores al 1,96. En otro caso, aún resultando orientativo de la tendencia, su interpretación no ofrece suficiente fiabilidad.

vez tomada en cuenta los efectos de las demás variables coligadas<sup>9</sup>. Al contrario, la no inclusión en la regresión de alguna de aquellas variables permitiría su libre variación al no controlar específicamente por el efecto de sus variaciones sobre la variable dependiente.

Así, la ecuación (1) mostraría el impacto de EM en el coste medio, dado el número de admisiones por cama. En efecto, CU como variable excluida debería variar en igual sentido que EM.

Si fuese CU la variable cuyo coeficiente interpretamos a partir de los resultados de la estimación, con FL incluida también en la regresión y siendo, sin embargo, EM la variable excluida, la interpretación sería otra. Puede derivarse de modo no ambiguo que un aumento de CU no iría en la dirección de una mayor eficiencia, ya que dicho incremento en la utilización vendría acompañado no de un aumento en el número de admisiones por cama (que se incluye en la regresión y por tanto se considera invariante), sino por el simple aumento de la estancia media (que es la que en la estimación debería variar, al no estar controlados sus cambios).

Otras combinaciones son también posibles con diferentes connotaciones en términos de mejora en la asignación de los recursos.

A la vista de distintas especificaciones de la función de costes hospitalarios estimada<sup>10</sup>, se pueden también analizar, (i) los diferentes resultados que se observan para los coeficientes asociados a CAS (esto es, el impacto de una variación en el margen en el número de enfermos tratados), y como éste depende, de nuevo, de cuál sea la variable de utilización incluida y cuál la excluida en la estimación, y (ii) las consecuencias que en términos de eficiencia (impacto en costes en relación a las variaciones de actividad) se derivan de dicho comportamiento<sup>11</sup>.

$$(2) \quad \begin{array}{l} \text{Costes totales} = 1.002,7 - 4.587,4 \text{ CU} + 91.778,1 \text{ EM} + \\ \text{(miles de ptas.)} \quad (2,97) \quad (-2,62) \quad (6,16) \\ \quad \quad \quad + 69,41 \text{ CAS} + \text{Vector de Combinación de casos} \\ \quad \quad \quad (12,43) \end{array}$$

$$R^2 = 0,88 \quad \bar{R}^2 = 0,85 \quad F(8,41) = 36,49$$

$$(3) \quad \begin{array}{l} \text{Costes totales} = 543,87 - 5.424,7 \text{ CU} + 77,93 \text{ CAS} + \\ \text{(miles de ptas.)} \quad (1,12) \quad (2,27) \quad (10,50) \\ \quad \quad \quad + \text{Vector Combinación de casos} \end{array}$$

$$R^2 = 0,76 \quad \bar{R}^2 = 0,72 \quad F(7,42) = 19,30$$

(9) Nótese que la inclusión de las tres variables en la regresión permitiría una interpretación ambigua, ya que el coeficiente aislado de cualquiera de ellas debería hacer referencia a su impacto parcial sobre la variable dependiente, permaneciendo constantes las otras dos variables, lo cual, en sentido estricto, no es posible entre EM, CU y FL.

(10) Estimaciones extraídas de «*A Budget-Bases Contract for Promoting Efficiency Improvements in the Hospital Sector*», G. López Casasnovas, 1984, Ph. D. Univ. de York (mimeo).

(11) Nótese, como antes, que las variables de utilización están vinculadas por la identidad:

$$\text{Estancia Media} = \frac{\text{Coeficiente de Utilización}}{\text{Admisiones por cama}} \times 365$$



output de la industria de cuidados de salud, y no del resultado de las intervenciones procedentes de los servicios sanitarios en cuanto a mejora de salud. La contribución de aquél a ésta se supone, aunque (i) no se conoce ni su carácter unívoco (existen intersectorialidades y tasas marginales de sustitución efectiva entre output de la industria sanitaria y otras intervenciones que no conviene ignorar), (ii) ni su correcta parametrización (¿se trata efectivamente de coeficientes fijos?), (iii) ni, por supuesto, su linealidad o «escala» (¿más actividad más resultado?, existiendo, cuando menos, rendimientos marginales decrecientes). En otras palabras, dicha aproximación no contempla procesos de sustitución de outputs para igual objetivo de maximización de la salud.

En primer lugar, por tanto, existe un problema de agregación del output, con un ajuste explícito tanto por casuística (gravedad de la demanda atendida) como por especialización (abanico de oferta de servicios puesta a disposición o «stand by»). No resolver correctamente dichos extremos condena a un determinado tipo de resultados: por ejemplo, es probable que la especialización, entendida como concentración de oferta, aminore los costes medios por la vía de la escala y concentración de los recursos, mientras que la complejidad, entendida como gravedad relativa de las patologías tratadas, los aumente. Dichos factores (complejidad y especialización) no se distribuyen homogéneamente entre los centros; en particular si el status de los centros difieren en cuanto a su propiedad (público, privado, con o sin ánimo de lucro) o a su naturaleza jurídica (fundaciones o entes públicos no diferenciados).

En segundo lugar, toda la utilización que la actividad recoge se supone necesaria: es decir, el output (intermedio) que se utiliza como referente de los costes medios, y de éstos para la eficiencia, se considera de idéntico valor sanitario. En otras palabras, las pautas de morbimortalidad se consideran no manipulables: todo lo hecho era necesario.

En tercer lugar, el análisis pivota en muchas ocasiones sobre la actividad, si nos referimos a centros hospitalarios de agudos, de internamiento. Sin embargo, éste no es el único output de una industria multiproducto como la hospitalaria, para la que cuentan también la consulta externa, las urgencias, las intervenciones ambulatorias y el seguimiento de determinados programas, con incidencia tanto en términos de actividad como de impacto en gasto. Acontece de modo parecido con la actividad de los centros de salud.

La cuestión del carácter de múltiples outputs de la actividad sanitaria hace que, de nuevo, se busque parametrizar la estimación, en caso de considerarse los costes medios por ingreso, o se proceda a su inclusión explícita como factor explicativo de alguno de dichos outputs adicionales. A esto último se resisten diversos trabajos, a efectos de no perder grados de libertad y ante la falta de series largas de datos; en particular, si se pretenden validaciones sobre la escala del output (con la inclusión de valores del output al cubo y al cuadrado) y sobre sus posibles complementariedades (a partir del signo negativo de los productos cruzados de los outputs). Como resultado de ello, algunos trabajos favorecen la utilización, directamente, de técnicas no paramétricas de estimación como el análisis envolvente de datos, que resultan más apropiadas para dicho tipo de problemas.

#### D. La valoración de la eficiencia de la institución sanitaria en abstracto

Una de las primeras distinciones que aparecen en la literatura comentada es la relativa a las estimaciones de coste, que reflejan los comportamientos observados, de aquellas otras que pretenden especificar una función de costes estrictamente dual a la función de producción. Desde esta segunda óptica, la cuantía del output y el precio de los inputs son las variables clave y no, por supuesto, detalles institucionales ajenos a la isocuanta de producción o a los precios relativos de los factores. Desde la primera, hace falta derivar la especificación de algún modelo concreto de comportamiento de las instituciones, otro que la simple maximización de beneficios, lo que no se suele producir en la realidad. Todo lo que no sea minimización de costes y distribución de excedente requiere algún supuesto acerca de lo que puedan ser (beneficiar) los ahorros internos en los costes o incrementos en output, ya que ambos factores están en la base de costes mínimos. Se trata, en todo caso, de estimar funciones de coste, no de ajustar con parsimonia ecuaciones de gasto. Para lo primero hace falta contar con una teoría previa de la que se derive algún tipo de optimización, mientras que para lo segundo predominarían factores puramente *ad hoc*.

#### E. Gasto versus coste de los servicios

En lo referente a la correcta cuantificación de los costes (imputación temporal) respecto de los gastos (salida de medios financieros), las dificultades se centran en las imputaciones temporales efectivas y, en definitiva, afectan al grado de homogeneidad con el que parten las estadísticas de base.

La contabilización temporal del gasto sufre el problema de que, en el sistema sanitario público, muchas veces lo presupuestado no es lo liquidado, ni lo liquidado, a veces, representa las obligaciones realmente contraídas, pagadas o no. Cuando se producen las afloraciones o deslizamientos, al incluir varios años, ello requiere una difícil imputación.

Otras cuestiones no menos importantes serían: a) la utilización indiscriminada de partidas relativas a costes de amortización que pueden responder a prácticas muy diversas, y b) la falta de preocupación por los índices de deflación de las cifras de gasto cuando la estimación temporal se realiza en términos reales.

#### F. Precio versus coste unitario de los factores

Hemos visto como, desde un punto de vista teórico al menos, una función de costes requiere introducir variables de precios de los inputs. Una primera alternativa es ignorarlos en base a suponer que los precios de los factores están dados y son idénticos sobre la geografía: así, por ejemplo, en el caso de los salarios del personal sanitario, publicados en el Boletín Oficial del Estado, o en el precio de los medicamentos. Este argumento confunde precio con coste unitario. En efecto, a precios idénticos los costes pueden diverger en razón a compensaciones no salariales o al disfrute de grados salariales superiores a los que se corresponden con las propias características del servicio. Pero si éste es el caso, no parece lógico que las estimaciones ajusten explícitamente por dichas diferencias, a no ser que se validen, previamente, como externos o exógenos a una correcta utilización de los recursos. Si éste no es el caso,

y la variable incluida es la de los costes observados, pese a que ello reporte una mejora del poder explicativo de la estimación, dicha alternativa no resultaría aceptable y, menos aún, bajo una supuesta lógica prescriptora de niveles de eficiencia económica.

De modo similar se podría aducir en lo referente a las políticas descentralizadas de compra de inputs cuando no existen precios que quepa considerar como externos o de mercado.

#### G. Referente de la noción de eficiencia

Dos son las posibilidades a dicho propósito: el considerar la eficiencia en base a la definición de «frontera» como mejor práctica observada, ya sea a partir de técnicas paramétricas o no paramétricas de estimación, o construir el referente de eficiencia a partir de la predicción muestral, sobre la base de las variables consideradas en el ajuste.

A su vez, la virtualidad de la utilización de uno u otro referente de eficiencia para objetivos de toma de decisión en materia de política presupuestaria resulta distinta, ya que, en el primero de los casos, se visualizan en mayor medida los errores de estimación («outliers») y los postulados efectuados en los procedimientos utilizados, mientras que en el segundo la idea de «promedio» facilita su aceptación.

#### H. Sobre las variables utilizadas y su definición

Algunos problemas adicionales a los ya vistos, y asociados con la definición de variables a efectos de estimación son:

(i) La utilización de «camas» como aproximación al capital, tamaño y dotación de los centros hospitalarios. En efecto, la consideración de la infraestructura como elemento fijo de las instalaciones hace que su inclusión o no en la estimación identifique a sus resultados como el horizonte temporal de análisis (el corto o el largo plazo). De modo similar en lo que se refiere a la interpretación de la constante como coste fijo de la función de costes.

(ii) La aproximación a la variable «docencia» en sus distintas alternativas, ya sea como variable dicotómica (0 si no docente, 1 de otro modo), continua (número de médicos en formación, por ejemplo) o a nivel de distinta submuestra (con especificación y formas funcionales diversas).

(iii) La consideración que deba merecer la investigación, ya sea en sus efectos sobre el output, a los costes unitarios o a la eficiencia en su conjunto.

#### I. Acerca de los métodos de estimación y elementos relacionados

Entre las técnicas de medición de la eficiencia (fronteras de producción) sobresalen las derivadas de la programación matemática y de los métodos econométricos<sup>12</sup>.

(12) Esta sección forma parte de un texto más amplio publicado por el autor en *Revista de Economía Aplicada*, 1995, con el título de «Análisis micro-micro de la eficiencia económica».

La aproximación econométrica es estocástica, y de este modo intenta distinguir los efectos del ruido estadístico de los efectos de la ineficiencia. Por el contrario, la programación es no-estocástica y confunde error e ineficiencia, llamando a su combinación ineficiencia.

La aproximación econométrica es paramétrica, y no distingue los efectos de errores de especificación de la forma funcional (tanto de la tecnología como de la ineficiencia) con la propia ineficiencia. La programación matemática es no paramétrica y menos propensa a este tipo de error de especificación.

Sería quizás por ello deseable hacer estocástica la programación así como más flexible la aproximación econométrica en lo que se refiere a su estructura paramétrica.

Veamos con un poco más de detalle cada una de las anteriores:

Por lo que se refiere a la *estimación econométrica* de la función de producción existen tres estrategias. La primera consiste en asumir que la ineficiencia es cero, con lo que, en ausencia de ineficiencia técnica, la estimación se simplifica possibilitándose la regresión mínimo cuadrática (MCO).

La segunda consiste en asumir que el error aleatorio es cero y estimar la función frontera *determinística*.

La tercera consiste en no asumir que dicho error es nulo y analizar la eficiencia técnica a través de estimar una función frontera *estocástica*. Ello permite descomponer el error en dos partes: un componente identifica la ineficiencia relativa de las observaciones respecto de la frontera eficiente (la cual presenta una distribución asimétrica), mientras que el otro componente incluiría los efectos de los errores de medición de las observaciones, variables no incluidas y *shocks* aleatorios. Su principal inconveniente, como veremos, radica en que no permite realizar la descomposición comentada de un modo individualizado, al referir las medidas de eficiencia únicamente al cálculo de las estimaciones medias, pudiéndose, en consecuencia, tan sólo estimar el nivel de eficiencia media para el total de la muestra. Para su estimación se utilizan tres técnicas distintas:

La *primera* se basaría en una corrección del método MCO, sin hacer hipótesis alguna acerca de la forma funcional del componente no positivo de la eficiencia. Se trata de estimar los parámetros por MCO y corregir el sesgo en el valor de la constante, dándole valores hasta que todos los residuos corregidos son no-positivos y al menos uno es cero.

La *segunda* técnica modificaría los MCO, estableciendo una determinada hipótesis sobre la forma funcional del componente no positivo de la eficiencia. La más comúnmente utilizada es la seminormal, que requiere la estimación de un parámetro adicional, la varianza de la distribución normal, que es truncada por encima de cero. Otra distribución de parámetro único es la exponencial. De dos parámetros serían la normal truncada y la gamma. Se trata, en definitiva, de estimar por MCO y modificar la constante así estimada aumentándola por el valor negativo de la media estimada de la ineficiencia, extraída de los momentos de los residuos MCO. Los residuos MCO se modifican en la dirección opuesta.



La *tercera* técnica es la del estimador de máxima verosimilitud (MLE). En su aplicación se asume una forma funcional para el componente no positivo de la eficiencia, y se estiman simultáneamente todos los parámetros tecnológicos de la función y los de la distribución del componente de la eficiencia.

Finalmente dos modelos de estimación se derivan de los *sistemas multiecuacionales y de los datos de panel*.

A pesar de que sea la función de producción *Cobb-Douglas* la más utilizada en este tipo de aproximaciones, las restricciones que contiene hacen que las formas funcionales flexibles sean mucho más interesantes. En este sentido, las especificaciones «translog» de la función frontera de producción tienen la ventaja de evitar o aminorar los problemas de colinealidad, en la medida en que se plantean como forma reducida de sistemas de ecuaciones simultáneas. El lema de Shephard permite utilizar ecuaciones definidas en términos de participaciones factoriales y un término de error.

El problema reside en cómo ligar este término de error con el propio de la translog (que es de coste) en relación al de las participaciones factoriales (que pueden responder a distintas lógicas). Para ello, los modelos transversales de ecuaciones múltiples resultan novedosos: dicha aproximación permite formas funcionales flexibles, acomoda la estimación de múltiples outputs e incorpora hipótesis de comportamiento de modo que resulta posible estimar no sólo la eficiencia técnica, sino también la económica o asignativa.

La estimación de *modelos de datos de panel* permite aprovechar la información que se tiene del productor en más de un período de tiempo. En la base de esta aproximación está la asociación entre efecto-específico del productor de la literatura de datos de panel con la literatura de ineficiencia de residuos de una sola cola de la literatura de funciones frontera.

El ruido estadístico varía aquí entre productores y a lo largo del tiempo, pero se puede asumir que la ineficiencia técnica varía sólo entre productores. Esta es obviamente una posibilidad. Las ineficiencias específicas a la empresa pueden juntarse con la constante y obtener un panel de datos convencional en el que no existe efecto temporal y el efecto productor es, en efecto, de una sola cola. Si la frontera estocástica se estimaba en un modelo longitudinal, se requerían hipótesis específicas sobre la forma funcional de los efectos específicos al productor y acerca de la independencia en la distribución de la ineficiencia respecto de los inputs: ello no es necesario si la información de corte transversal se combina con los de corte temporal como ocurre en los datos de panel.

A efectos de estimación tres son los métodos disponibles: el «intra estimador» (cada ecuación transformada en desviaciones de las mismas variables en el tiempo para cada productor) no postula sobre la distribución del componente de ineficiencia del residuo ni sobre su independencia de los inputs. El estimador de Mínimos Cuadrados Generalizados no postula nada sobre lo primero, pero sí sobre lo segundo, al tener que permitir la inclusión de regresores en el modelo que no varían en el tiempo, y el estimador de MLE requiere las dos hipótesis anteriores, tanto la relativa a la distribución como la referida a la independencia de la ineficiencia de los inputs productivos.

La estimación por máxima verosimilitud ofrece, en este sentido, una aproximación superior, al permitir conocer la estructura del error compuesto (papel del factor aleatorio y del residuo de ineficiencia, una vez asumida la independencia de ambos componentes del error compuesto).

En definitiva, en la estimación basada en datos de panel, la consideración de efecto fijo permite una constante distinta para cada unidad de análisis. Normalizada a valor unitario la que tiene el valor más alto, ello permite un ranking de indiciación de la eficiencia relativa. El carácter invariable en el tiempo, sin embargo, en presencia de variaciones en la tecnología, cuestiona este tipo de aproximación.

La consideración de efecto aleatorio identifica, de otro modo, que constituye ineficiencia lo que puede ser un simple cambio de tecnología. Dicha crítica se aminoraría en el caso de que el panel se refiriera a una serie temporal relativamente corta.

Los modelos de *programación matemática* en su versión Análisis Envolvente de Datos (DEA) tienen a su favor que miden la eficiencia técnica, con lo que no se consideran los precios (cuestión ésta compleja en toda comparación público-privada), y hace innecesario el ponderar variables de modo previo al análisis. La eficiencia técnica se convierte, de este modo, en un terreno aséptico para la comparación de la eficiencia relativa entre el sector público y privado. La actividad realizada se mide, de este modo, en términos de la capacidad del productor de expansionar su vector de outputs dadas las restricciones impuestas por las mejores prácticas observadas. El índice de eficiencia es el ratio en el que todos los inputs podrían reducirse, manteniendo los outputs constantes, si el productor realizara su actividad como lo hace el productor hipotético (combinación lineal de las dos observaciones de mejor práctica). En cualquier caso, la eficiencia plena exigiría que no hubiera *slack* ya sea en inputs o en outputs, ya que su corrección también permitiría mejorar la actividad del productor.

La aplicación del *Análisis Envolvente de Datos* en el ámbito público puede ser particularmente deseable al acoger vectores multidimensionales de output. A diferencia de la mayoría de métodos paramétricos (excepto los que utilizan datos de panel), ofrece una información singularizada a nivel de cada productor de la eficiencia. Como vimos, no utiliza información sobre precios, ni ponderaciones más o menos arbitrarias para agregar los ratios de eficiencia.

La conveniencia de introducir mayor flexibilidad en las técnicas de programación se presenta sin embargo hoy como un *desideratum* para la mayoría de investigadores en esta área. Algunas de las modificaciones en los modelos DEA tradicionales vienen de la mano de:

- a) Multiplicadores limitados, que en el campo de la producción vienen a dar una forma más plana a la isocuanta base del análisis de eficiencia radial.
- b) Debilitar la hipótesis de convexidad de la función de producción (por ejemplo, las combinaciones entre observaciones no necesariamente han de existir: la tecnología puede no permitir combinaciones lineales sino escalonadas).

- c) Relajar la hipótesis de monotonidad, o supuesto de disponibilidad fuerte; por ejemplo, en el análisis de la función de producción, dificultando o haciendo imposible crecer y contaminar a la vez. El coste de uno en relación al otro puede, además, analizarse como dual de costes (precio sombra de la contaminación en términos de output).

Y, contrariamente, en lo que se refiere a la elevada flexibilidad de las ponderaciones que contiene el DEA, se pretende más bien en la actualidad restringirlas, a efectos, por ejemplo, que no permita ponderaciones nulas de inputs o outputs que se consideran relevantes, o que dichas ponderaciones sean diferentes entre las unidades analizadas<sup>13</sup>.

En resumen, se estaría produciendo, por una parte, un intento de flexibilizar algunos aspectos del DEA (por ejemplo, en el caso de rendimientos de escala variables) en relación a lo que han sido los modelos no paramétricos convencionales (convexidad, rendimientos de escala constantes) y, por otra parte se estaría introduciendo una cierta «parametrización» (restringiendo, por ejemplo, la libre disponibilidad de inputs y outputs, el valor de las ponderaciones, etc.). Ello contrasta con lo que parece acontecer con las técnicas paramétricas, para las que se observa la tendencia a no aceptar plenamente la estocasticidad, restringiendo la forma funcional o utilizando procedimientos de estimación corregidos o modificados<sup>14</sup>.

#### 4. ALGUNOS DE LOS ANÁLISIS POSIBLES Y SUS IMPLICACIONES PARA LA GESTIÓN

##### 4.1. Economías de escala y de alcance

Las consideraciones relativas a la función de costes hospitalarios, relevantes a la vista de la estructura de costes y de la gama de productos sanitarios suministrados desde los centros hospitalarios, serían las siguientes:

1. Reconocimiento del carácter multiproducto de la actividad de los centros, y la posible existencia de economías de gama o alcance.

Se entiende existen economías de gama si el output global de una organización (multiproducto) es mayor que el output que podrían obtener diversas organizaciones (uniproducto) que, entre todas ellas, empleasen la misma cantidad de inputs.

La implicación de ello sería que los costes de producción disminuyen al producir de modo conjunto los distintos productos en una misma organización, en lugar de especializarse.

2. Estudio de los costes incrementales y los costes incrementales medios existentes entre los distintos outputs, asociando su comportamiento a la posible presencia de economías de escala.

(13) Véase al respecto, Pedraja Fco., Salinas J. y Smith P. (1994) «*La restricción de las ponderaciones en el análisis envolvente de datos: Una fórmula para mejorar la evaluación de la eficiencia*», Univ. de Extremadura (mimeo).

(14) Una descripción de métodos y alternativas puede leerse en el trabajo de G. López y A. Wagsstaff incluido en *La regulación de los servicios sanitarios en España*, G. López y D. Rodríguez (coords.), Ed. Cívitas, Madrid, 1997.

El concepto de coste incremental indicaría, en el contexto de una función de producción multiproducto, la adición a los costes que resulta de incrementar en una unidad el nivel corriente del output en cuestión «ceteris paribus» el resto de actividad. El coste incremental se definiría como:

$$IC_1 = C(Y_1, Y_2) - C(0, Y_2)$$

A su vez, el coste incremental medio AIC indicaría el coste extra asociado a la producción del output en cuestión, promediado por encima de la cantidad ya producida de dicho output<sup>15</sup>. Es decir,

$$AIC_1 = IC_1 / Y_1$$

Nótese, por tanto, que AIC es el coste medio en el caso de la producción de un sólo output. Si AIC registra una forma-U, con AIC por encima del coste marginal, nos situaríamos a la izquierda del mínimo de la curva AIC, de modo que existirían economías de escala específicas de producto; mientras que de-seconomías de escala existirían si los costes medios incrementales fueran inferiores a los costes marginales.

En este sentido, el análisis de la existencia de economías de escala se ha de distinguir del estudio del grado de sobrecapitalización que pueda tener un centro hospitalario. Mientras el primero se refiere al efecto en los costes de expansiones sucesivas en output; es decir, se analiza, para detectar la presencia de economías de escala, el comportamiento de los costes sanitarios si los niveles de output son más o menos elevados, para la sobrecapitalización se analiza, al contrario, si el stock de capital es o no demasiado grande, dado, precisamente, un nivel de output determinado.

En cualquier caso, los dos aspectos anteriormente reseñados son importantes si se desean considerar los incentivos que generan en las decisiones sobre la producción hospitalaria distintos métodos de financiación. Nótese, por ejemplo, que las relaciones entre costes marginales y medios de cada uno de los productos hospitalarios, en relación a los precios o tarifas de la actividad financiada, pueden afectar las decisiones de expandir o reducir una determinada actividad, alterar la combinación de servicios clínicos o reposicionar el hospital en un entorno que pueda resultar más competitivo<sup>16</sup>.

#### 4.2. Productividades y retribuciones factoriales relativas

Nos referimos aquí a la utilización de las estimaciones de la función de producción, o su dual de costes como mecanismo con los siguientes rasgos:

El output  $Y$  depende de un vector de inputs  $X_i$ :  $Y=f(X_i)$

De su estimación se deriva el coeficiente de productividades marginales de dos factores productivos cualquiera:  $f_i(\cdot)/f_j(\cdot)$ , que se compara con el ratio de precios o costes relativos de dichos inputs:  $C_i/C_j$ .

(15) Nótese, por tanto, que el ratio entre ambos equivale a economías de escala específicas al producto.

(16) En la medida que se liberalizasen en gran medida los flujos sanitarios, ello podría plantear un problema de endogeneidad del output (al poder elegir el hospital los niveles de output según sus costes relativos), que haga que los coeficientes estimados en una regresión múltiple respecto de la actividad, resulten sesgados.

Para el caso más simple de una función Cobb-Douglas<sup>17</sup>, de su estimación log lineal resulta posible derivar los valores de las elasticidades de los inputs  $B_i$ . Para esta función, el coeficiente de las elasticidades entre dos inputs cualesquiera ( $B_i/B_j$ ) ha de igualar el nivel de gasto relativo de dichos inputs

$$[(X_i)(C_i)]/[(X_j)(C_j)]$$

Cualquier desviación observada entre los dos cocientes indica la pauta de corrección apropiada, que se habrá de expresar en variaciones de  $C$ ,  $X$ , o en combinaciones de *ambas*.

La anterior aproximación requiere la estimación, a través de técnicas econométricas de distinta complejidad, de la función de producción que permite, en definitiva, separar las cuestiones de productividad y de coste al considerar la eficiencia económica en el gasto.

#### 4.3. Limitaciones del análisis en su aplicación a la gestión

Quisiéramos cerrar este trabajo refiriéndonos a la controversia sobre la posibilidad real de insertar los anteriores resultados en una realidad marcada por un sistema determinado de presupuestación de instituciones y de retribución de profesionales. En ambos sistemas, como es conocido, predominan las rigideces en los ajustes. Ello se traduce en lo que se pueda considerar, en cada caso concreto, «coste eludible», a efectos de la actuación pública al servicio de la eficiencia económica. Este es el caso, por ejemplo, de aquellos supuestos en los que existe un exceso de capacidad, o de aquellos servicios públicos en los que el carácter multi-producto de la actividad, generador de economías de alcance o gama, dificulta en la práctica el análisis de eficiencia de un modo aislado (nótese que los costes unitarios dependen tanto de la escala de actividad como de la gama de servicios suministrados). Así, la valoración de lo que constituye propiamente un coste fijo, semi-fijo o variable, a efectos de consideración de costes unitarios, tiene mucho que ver con el tipo de contrato vigente para los distintos factores de producción, y de la posibilidad, cuando menos, de reubicar los factores en el seno de las instituciones, con costes nulos. Ninguna de estas dos variables es, en todos los casos, endógena a la gestión y, por tanto, no cabe siempre ignorarla a efectos de financiación de compra eficiente de servicios. El optimismo sobre las reformas no puede convertirse en este caso en ingenuidad.

## REFERENCIAS BIBLIOGRÁFICAS

- Aigner, D. J.; Lovell, C. A. K. y Schmidt, P. (1977): «Formulation and estimation of stochastic production function models», *Journal of Econometrics*, 6, págs. 21-37.
- Arellano, M. y Bover, O. (1990): «La econometría con datos de panel», *Investigaciones Económicas*, (2.<sup>a</sup> época) XIV, págs. 3-45.

(17) Para funciones más complejas, tales como la función de forma flexible «translog», el coeficiente de productividades marginales, que se deriva de la correspondiente estimación, se ha de comparar a los costes relativos de los inputs. En este caso, la elasticidad de cada uno de los inputs depende, a diferencia de lo observado en la función Cobb-Douglas, de la cantidad también utilizada de los otros inputs.

- Barer, M. L. (1982): «Case mix adjustment in hospital cost analysis: Information Theory revisited», *Journal of Health Economics*, 1, n.º 1, págs. 53-80.
- Barlow, R. (1968): «Review of Economic Analysis for Health Service Efficiency», *Economic Journal*, 78, págs. 921-923.
- Breyer, F. (1987): «The specification of a hospital cost function. A comment on the recent literature», *Journal of Health Economics*, 6, págs. 147-157.
- De Fraja, G. (1993): «Productive efficiency in public and private firms», *Journal of Public Economics*, 50.
- Feldstein, M. S. (1967): *Economic Analysis for Health Service Efficiency: Econometric Studies of the British National Health Service*, North-Holland, Amsterdam.
- Forsund, F. R.; Lovell, C. A. K. y Schmidt, P. (1980): «A survey of frontier production functions and their relationship to efficiency measurement», *Journal of Econometrics*, 13, págs. 5-25.
- Fried, H. O.; Lovell, C. A. K. y Schimdt, S. S. (eds.) (1993): *The measurement of productive efficiency. Techniques and applications*, Oxford Univ. Press.
- Gaynor, M. y Anderson, G. F. (1995): «Uncertainty demand and structure of hospital costs, and the cost of empty beds», *Journal of Health Economics*, 14, págs. 291-317.
- González López, B. y Barber, P. (1996): «Changes in public Spanish hospital efficiency after the program contracts», *Investigaciones Económicas*, v XX, 3.
- Good (1992): «Productive efficiency and contract management: some evidence from public transit agencies», *Public Finance Quarterly*, abril, págs. 195-215.
- Grannemann, T. W.; Brown, R. S. y Pauly, M. V. (1986): «Estimating hospital costs: A multiple output analysis», *Journal of Health Economics*, 5, págs. 107-127.
- Grosskopf, S. y Valdmanis, V. (1987): «Measuring hospital performance: a non-parametric approach», *Journal of Health Economics* 6, págs. 89-108.
- Lovell, C. A. K. (1995): «Econometric Efficiency Analysis: A policy oriented review», *European Journal of Operational Research*, 80, págs. 452-461.
- Laffont, J. J. y Tirole, J. (1993): *A Theory of Incentives in Procurement and Regulation*, MIT. Press.
- López-Casasnovas, G. (1995): «Cost efficiency estimation and rate setting in Health Care», *Third European Conference on Health Economics*, Stockholm, mimeo.
- López-Casasnovas, G. y Wagstaff, A. (1993): «Competencia en Sanidad», *Moneda y Crédito*, 196, págs. 181-229.
- Newhouse, J. P. (1994a): «Reimbursement under uncertainty: What to do if one cannot identify an efficient hospital», *Santa Mónica, RAND MR-309-hcfa*.

- Newhouse, J. P. (1994b): «Frontier Estimation: How useful a tool for Health Economists». *Journal of Health Economics*, 13, págs. 317-322.
- Pope, G. C. (1990): «Using Hospital-Specific Costs to improve the Fairness of Prospective Reimbursement», *Journal of Health Economics*, 9, págs. 237-251.
- Reichelstein, S. (1992): «Constructing Incentive Schemes for Government Contracts: An application of the Agency Theory», *The Accounting Review*, vol. 67, 4 Oct., págs. 712-731.
- Shleifer, A. (1985): «A Theory of Yardstick Competition», *RAND Journal of Economics*, 16, págs. 319-327.
- Schmidt, P. y Lovell, C. A. K. (1979): «Estimating technical and allocative efficiency relative to stochastic production and cost frontiers», *Journal of Econometrics*, 9, págs. 343-366.
- Schmidt, P. y Sickles, R. C. (1984): «Production frontiers and panel data», *Journal of Business and Economic Statistics*, 2, págs. 367-374.
- Wagstaff, A. (1987): «Measuring technical efficiency in the National Health Service: a stochastic frontier analysis», *Discussion Paper*, 30, Centre for Health Economics, University of York.
- Wagstaff, A. (1989b): «Estimating efficiency in the hospital sector: a comparison of three statistical cost frontier models», *Applied Economics*, 21, págs. 659-672.
- Wagstaff, A. y López, G. (1995): «The Measurement of Hospital Efficiency: A Comparison of Alternative Stochastic Frontier Models», *Third European Conference on Health Economics, Estocolmo*, mimeo.

#### ABSTRACT

This article deals with some recent developments in efficiency analysis in health care services and their application, especially in the area of budgeting.

Alternative budgetary practices are proposed based on efficiency analysis and prospective budgeting, leading to an analysis of the contributions made by health economics to the study of efficiency in the provisioning of health care services and to the problems and methods which are raised in this area. The paper concludes by linking aspects related to efficiency analysis with budgeting concerns. Likewise, some of the limitations of the analysis, as applied to management, are pointed out.

*Key words:* efficiency, health care services, budgeting.