

R. Fare, S. Grosskopf y C.A.K. Lovell
Production frontiers

Cambridge, Cambridge University Press, 1994

H.O. Fried, C.A.K. Lovell y S. S. Schimdt (eds.)
**The measurement of productive
efficiency. Techniques and applications**

Oxford, Oxford University Press, 1993

Z. Griliches (ed.)
**Output measurement in the
services sector**

Chicago, NBER, 1992

Z. Griliches y J. Mairesse (eds.)
**Productivity issues in the services at the
micro level**

London, Kluwer Academic Publishers, 1993

GUILLEM LÓPEZ CASASNOVAS
Universidad Pompeu Fabra

Resulta evidente que el debate sobre la eficiencia relativa público-privada requiere de un instrumental de análisis mucho más preciso que el que permite la simple comparación de costes o de niveles de actividad entre ambos sectores. El carácter controvertido de la aproximación o las propias ambigüedades en la interpretación de los resultados, aun aplicando técnicas relativamente sofisticadas, son comunes en el ámbito referido¹. Además, anotemos en todo caso que, para algunos, siempre cabrá interpretar que lo que acaba legitimando la provisión pública no es su eficiencia relativa sino algún otro objetivo cuyo grado de cumplimiento no se recoge en el tipo de evaluación anterior.

Por ello, el análisis de la *eficiencia 'intra' sector público*, o eficiencia relativa entre unidades suministradoras de los servicios, supone un campo mucho más acotado. Las responsabilidades asumidas y los objetivos fijados para las unidades suministradoras de los servicios —y aun en el caso de que exista cierta ambigüedad

(1) Véase a este respecto el trabajo de Ley (1991) y la crítica anexa a su trabajo realizada por J. Ginestal.

en su definición— debieran ser, en principio, idénticos en relación a todos y cada uno de los ciudadanos, a la vista al menos del ordenamiento público y del estatuto de las agencias proveedoras de los servicios.

En razón a lo anterior, empieza a sobresalir ya, también en nuestro país, una línea de investigación que busca un análisis 'micro- micro' de la eficiencia pública; es decir, un análisis referido al ámbito interno de sector público, y centrando el foco de atención en un servicio concreto.

Son aportaciones recientes a dicha literatura, en el campo no paramétrico, los trabajos de I. Vilardell y D. Prior y M. Solá, del Departamento de Economía de la Empresa de la Universidad Autónoma de Barcelona, con la aplicación del Análisis Envolvente de Datos (DEA) a un conjunto de servicios municipales en el primer caso² y a los hospitales catalanes financiados publicamente, en el segundo³. A. Cuenca (1993), de la Universidad de Zaragoza, ha aplicado idéntica técnica en el análisis de la eficiencia de la provisión de servicios de extinción de incendios, utilizando datos de la Dirección General de Protección Civil, y F. Pedraja y otros, de la Universidad de Extremadura, en el análisis de la administración de la Justicia.

En el ámbito paramétrico, a los trabajos iniciales de López (1985) y López y Wagstaff (1993) se añaden los de Puig (1992), González y Villalobos (1993), Leote (1992) y Quintana (1994), entre otros.

Una parada y fonda de recapitulación del camino recorrido y la presentación de cuatro textos recientes que pueden dar un nuevo empuje a futuras investigaciones es el objetivo de las líneas que siguen.

1. TÉCNICAS AL SERVICIO DEL ANÁLISIS MICRO-MICRO DE LA EFICIENCIA ECONÓMICA

Los textos de referencia que sirven al comentario de esta primera sección son los de Fare, Grosskopf y Lovell (1994) y Fried, Lovell y Schimdt (1993). Las nuevas aportaciones en el esfuerzo de medir la eficiencia económica permiten discutir los puntos fuertes y débiles de las técnicas de medición de la eficiencia productiva.

Tres son sus acepciones:

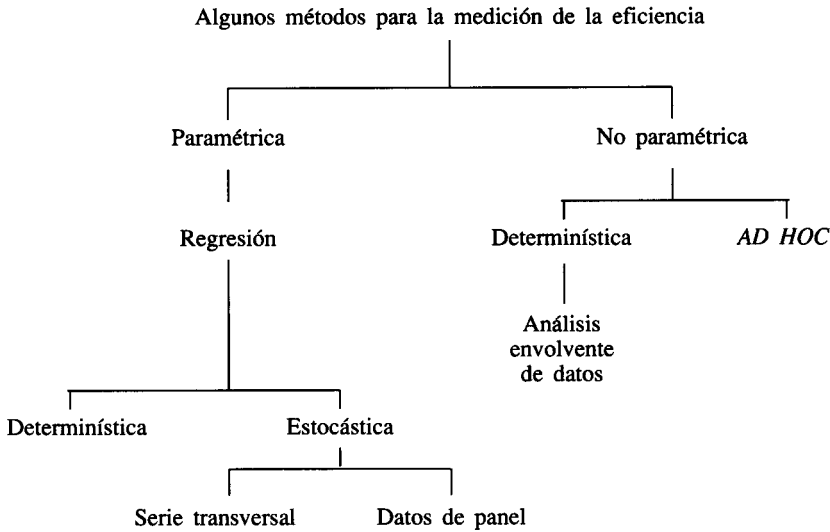
1. Técnicas paramétricas ('Goal programming' y aproximaciones econométricas).
2. Técnicas no paramétricas y no estocásticas (DEA).
3. Técnicas no paramétricas y estocásticas (DEA restringido).

En efecto, entre las técnicas de medición de la eficiencia (fronteras de producción) sobresalen las derivadas de la programación matemática y de los métodos econométricos.

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.

(2) Véase I. Vilardell (1990).

(3) D. Prior y M. Solá (1993). Una aplicación similar, relativa a la Asistencia primaria, está siendo iniciada en los Departamentos de Economía de la Empresa de las Universidades de Oviedo y de Zaragoza.



Fuente: Adaptado de M. Barrow y A. Wagstaff (1989).

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, posibilitándose 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 incluídas 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, estimar tan sólo 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 semi-normal, 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.

Cabe destacar que, en cualquier caso, los *rankings* que resultan de los MCO corregidos y de los modificados son idénticos a los derivados por MCO; sólo a efectos de considerar magnitudes se justificaría la utilización de los primeros.

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 hace 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 una panel de datos convencional en el que no existe efecto temporal y el efecto productor es 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 la 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 indicación 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* dada las restricciones impuestas por las mejores prácticas observadas. El índice de eficiencia es la *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 las *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 monotonicidad, 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, de que no permita ponderaciones nulas de *inputs* o *outputs* que se consideran relevantes, o que dichas ponderaciones sean diferentes entre las unidades analizadas⁴.

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 modificados⁵.

2. LA MEDICIÓN DEL *OUTPUT* DE LAS ACTIVIDADES PÚBLICAS OBJETO DE ANÁLISIS MICRO-MICRO DE LA EFICIENCIA PÚBLICA

Los textos de referencia, que sirven al comentario de esta segunda sección son: Griliches (1992) y Griliches Z y J, Mairesse (1993). [Algo más antiguos, pero también de interés son los textos de: Cave, Kogan y Smith (1990) y Bentkover y otros (1986)].

Para la mejora de la gestión pública –y, obviamente, para la utilización de las técnicas anteriores– resulta ineludible la cuantificación del *output* de la actividad pública.

Una primera aproximación a la medición del *output* resultante de las intervenciones del sector público requiere distinguir –a efectos de su posterior evaluación– entre aquellas actividades en las que tiene un importante protagonismo la demanda proveniente de los propios usuarios, respecto de aquellas otras actuaciones para las que el sector público es un simple ofertante de determinados servicios. Y éstas últimas, de aquéllas otras actividades en las que la oferta, estimada colec-

(4) Véase al respecto F. Pedraja, J. Salinas y P. Smith (1994).

(5) Nótese que ello supone asumir que la estructura de la función de producción es la misma entre eficientes e ineficientes cuando, en realidad, puede ser precisamente que los más eficientes tengan una estructura distinta al resto (por ejemplo, explotando las elasticidades de sustitución de manera diversa). Estas restricciones son criticadas por Fare, Grosskopf y Lovell (ob. cit.).

tivamente como necesaria, se impone al individuo, incluso pese a su posible falta de interés (véase cuadro adjunto).

A su vez, de entre las actuaciones más vinculadas a las preferencias individuales, la demanda puede tener un carácter regular o recurrente, o esporádica. La regular puede venir medida en términos de volumen –si todos los casos pueden considerarse como similares– o puede tener carácter específico, como actividad relativa a casos que son singulares en su especificidad. Lo anterior se contrapone a la naturaleza de lo que puede ser la actividad derivada de una demanda esporádica, no ‘almacenable’ en su provisión, y no planificable con certeza en la provisión acerca de su utilización efectiva a lo largo del tiempo (proyectos singulares, de desarrollo incierto, intervenciones tan urgentes como impredecibles, etc.).

Para las actividades públicas que no son secuencia inmediata de la demanda (actividades regulatorias, preventivas, de mantenimiento de servicios generales, etc.), la lógica es muy distinta: el éxito puede consistir, precisamente, en que no se haga efectiva la utilización del servicio que se previene, que no sea necesaria la ‘reposición’ –consumo del servicio–, debido a los efectos nocivos que, de otro modo, se producirían. Su medida ha de fijarse en términos distintos a los de consumo efectivo, como, por ejemplo, a través del cumplimiento de determinados estándares, de la reducción de ciertas desviaciones no deseadas, o en medidas de satisfacción de los usuarios.

Ciertamente, los *outputs* públicos incluyen elementos cualitativos y no sólo cuantitativos; de impacto en el bienestar de los usuarios y no sólo de producción de actividad. Sin embargo, sin un esfuerzo en mejorar el conocimiento de lo segundo mucho más difícil se nos antoja lo primero.

El cuadro de la página siguiente, muestra las alternativas existentes en cada disyuntiva para la medición de los *outputs* públicos.

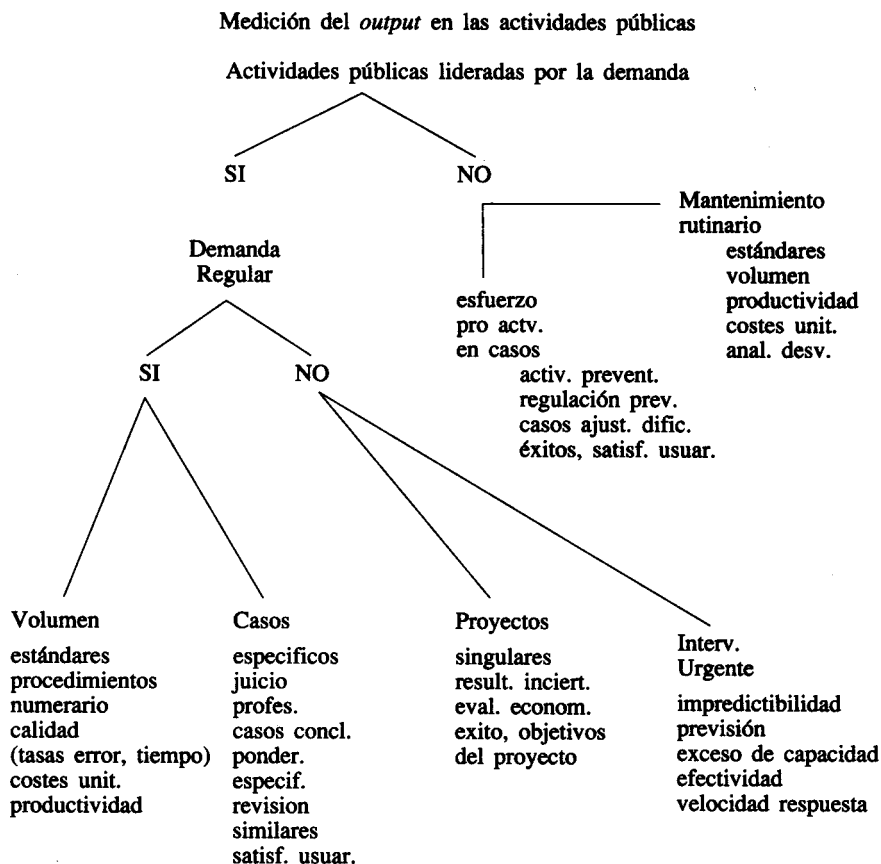
Los dos libros que encabezan esta sección vienen, en el contexto anterior, a suministrar más aplicaciones en el ámbito genérico del análisis de la eficiencia de los servicios, y de los servicios públicos en particular: por ejemplo, en el ámbito de los centros asistenciales de cuidados ambulatorios, servicios de transporte, de aseguramiento, educación y de determinados servicios asistenciales. Los textos ayudan por lo demás a entender el tratamiento que otorga la Contabilidad Nacional a la productividad de los servicios y las ventajas e inconvenientes de algunas técnicas al uso para la medida de cuestiones como la eficiencia técnica, las economías de escala y alcance o la calidad en los servicios personales.

En definitiva, los libros referidos refuerzan, a nuestro entender, el argumento principal de esta recensión: el interés de una aproximación micro-micro de la eficiencia económica; aproximación que pasa por la necesidad de cuantificar el *output* de los servicios y asociar a dichos *outputs* la tecnología capaz de derivar los *locus* de eficiencia en la evaluación de la actividad.

Para ello, las técnicas referidas en la primera parte del presente texto constituyen elementos sustantivos del estado de conocimiento sobre dicho ámbito de análisis de la gestión pública, que permiten aplicaciones inmediatas para el cálculo de la eficiencia.

Podrá pensar algún lector que ésta es una visión muy optimista y excesivamente tecnicista de la aproximación a la eficiencia pública. Pero dichas reticencias, a mi entender, han de perder peso si se analiza la aproximación propuesta –a través de la bibliografía reseñada– a la luz de las alternativas existentes.





Fuente: Lewis y Jones (1990).

REFERENCIAS BIBLIOGRÁFICAS

- Barrow, M. y A. Wagstaff (1989): "Efficiency measurement in the Public Sector: an appraisal", *Fiscal Studies*, vol. 10, págs. 73-97.
- Bentkover y otros (1986): *Benefits assesment: the state of the art*, D. Reidel publisher, Boston.
- Cave, M.; M. Kogan y R. Smith (1990): *Output and performance measurement in Government: the state of the art*, Jessica Kingsley, London, págs. 124-142.
- Cuenca, A. (1993): *La medición de la eficiencia en el sector público. Una aplicación a los servicios de protección contra incendios*. Tesis doctoral no publicada, Universidad de Zaragoza.

- González-Valcárcel, B. y J. Villalobos (1993): "Indicadores de actividad y costes en hospitales españoles", *Hacienda Pública Española*, Monografías, n.º 1, 1993, págs. 127-141.
- Griliches, Z. (1992): *Output measurement in the service sector*, The University of Chicago Press: National Bureau of Economic Research.
- Leote, R. (1992): *La medición de la eficiencia en el sector hospitalario. El caso portugués*, Tesis doctoral, Universidad Técnica de Lisboa.
- Lewis y Jones (1990): "Output and performance measurement in Government. The state of the art", en M. Cave *et al.* (eds.), J. Kingsley publishers, Londres.
- Ley, E. (1991): "Eficiencia productiva: un estudio aplicado al sector hospitalario", *Investigaciones Económicas*, vol. 15, n.º 1, págs. 71-88.
- López Casanovas, G. (1985): "Gasto público y racionalización presupuestaria", *Papeles de Economía Española*, n.º 23, págs. 203-213.
- López Casanovas, G. y A. Wagstaff (1993): "Eficiencia y competitividad en los servicios públicos: algunas consideraciones relativas a la asistencia sanitaria", *Moneda y crédito*, n.º 196, págs. 181-232.
- Pedraja, F.; J. Salinas y P. Smith (1994): *La restricción de las ponderaciones en el análisis envolvente de datos: una fórmula para mejorar la evaluación de la eficiencia*, Universidad de Extremadura, mimeo (pendiente de publicación).
- Puig Junoy, J. (1991): *La provisió de serveis hospitalaris. Creixement, inflació i tecnologia en el sector hospitalari espanyol: 1974-1985*, Tesis doctoral, Universidad Autónoma de Barcelona.
- Prior Jiménez, D. y M. Solá i Tey (1993): *L'eficiència dels hospitals de Catalunya. Comparació entre els hospitals públics i els privats*, Barcelona, Generalitat de Catalunya, Departament de Sanitat i Seguretat Social.
- Quintana, J. (1994): *Eficiencia relativa en la red de hospitales públicos españoles*, Fundación BBV, (pendiente de publicación).
- Vilardell, I. (1989): "La evaluación de la eficiencia de las actividades de la Administración pública", *De Economía Pública*, vol. 2, n.º 1, págs. 47-60.