

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 erreparatu dira: jardueren input eta output direlakoan 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.

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, exce-

lencia, 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 (*Decision 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 Barcelona³. 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 atribuyen⁴. 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.

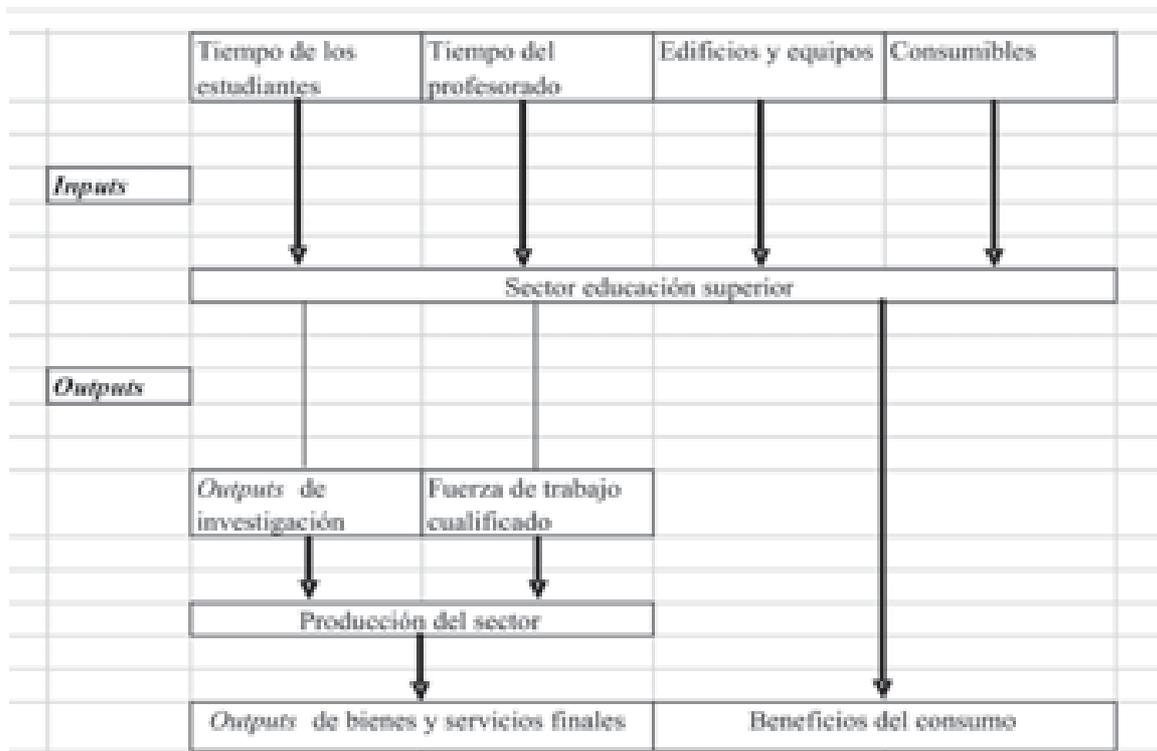
Como en él se ve, las universidades utilizan un conjunto de recursos (tiempo de los estudiantes, tiempo del profesorado, consumibles y equipos y edificios) para producir una serie de *outputs* que pueden ser clasificados, en términos generales, en productos de la docencia (conocimientos adquiridos por 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 servicios⁵. 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 N°. 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 n° 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 compro-

bará 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 comprensión 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 superior⁶.

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 aquellas 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 efi-

ciencia 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 cuestionable⁸.

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.

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 repre-

senta 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 universidades⁹.

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 indisoluble 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 universidades¹⁰.

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 bibliotecónica, 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 *Citation 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:

$$F1 (\text{año } t) = \frac{\text{Citas recibidas en } t \text{ por los artículos publicados } n \text{ años antes}}{\text{Artículos publicados en } m \text{ años antes}}$$

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 ellas¹². 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).

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 ineficientes¹⁴.

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 incluir¹⁵.

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:

$$\text{Max } \omega_0 = \theta_0^A \omega_0^A + \theta_0^B \omega_0^B$$

sujeto a:

$$\sum_{j=1}^n \lambda_j^A x_{Aj} \leq x_{A0}$$

$$\sum_{j=1}^n \lambda_j^B x_{Bj} \leq x_{B0}$$

$$\sum_{j=1}^n \beta \lambda_j^A x_{Cj} + \sum_{j=1}^n (1-\beta) \lambda_j^B x_{Cj} \leq x_{C0}$$

$$\sum_{j=1}^n \lambda_j^A y_{Aj} \geq \omega_0^A y_{A0}$$

$$\sum_{j=1}^n \lambda_j^B y_{Bj} \geq \omega_0^B y_{B0}$$

$$\lambda_j^A, \lambda_j^B \geq 0$$

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 Casanovas, 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 canadienses¹⁷.

En la mayoría de los trabajos anteriores, aparte de realizar las segmentaciones de la muestra recién mencionadas, los auto-

res 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 medicina¹⁸.

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 Casanovas, 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 las 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 activi-

dades 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.

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 conti-

nuas 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 <i>senior colleges</i> y universidades en el estado de Texas en el periodo 1981-85. 161 <i>Doctoral institutions</i> 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 <i>Doctoral institutions</i> 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 <i>Doctoral institutions</i> 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 <i>Doctoral institutions</i> 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 <i>research universities</i> (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 <i>research universities</i> .	CCR.

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 <i>The Times Good University Guide</i> 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.

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. <i>Value of non-current assets</i> .	Número de matriculados (ETC). Matriculados en cursos para no graduados. Matriculados en cursos para graduados. Títulos de grado concedidos. Títulos de postgrado concedidos. <i>Research Quantum</i> (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 (<i>South African Post-Secundary Education</i>).	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.

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.

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.

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 DOUCOLIAGOS, 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 Sciences*, 22, 6, págs. 253-257.
- ALBI, E.; GONZÁLEZ-PÁRAMO, J. M. y LÓLEZ CASASNOVAS, G. (1997): *Gestión Pública. Fundamentos, técnicas y casos.*, Barcelona, Ariel.

- ;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 Sciences*, 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 Autònoma 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 LEOPOLDSEDER, 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 Measurement 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.