

# **EFICIENCIA EN EDUCACIÓN SECUNDARIA Y FACTORES EXÓGENOS: SENSIBILIDAD DE LOS RESULTADOS ANTE MODELOS ALTERNATIVOS**

Cordero Ferrera, José Manuel

Pedraja Chaparro, Francisco

Salinas Jiménez, Javier

*(Universidad de Extremadura)*

*(Instituto de Estudios Fiscales)*

## **ABSTRACT**

El objetivo de este trabajo es estudiar la sensibilidad de los índices de eficiencia estimados para los distintos centros educativos a las diferentes alternativas propuestas en la literatura para incluir los factores exógenos en el análisis. Concretamente, nos centraremos en dos cuestiones fundamentales. Contando con una información muy rica suministrada por unas encuestas sobre estas variables consideraremos, por un lado, la posibilidad de incluir directamente las variables más relevantes en el análisis de eficiencia o de utilizar, alternativamente, el análisis de componentes principales para sintetizar la información recogida en dichas variables. Por otro lado, la comparación entre el enfoque que considera que estos factores deben tenerse en cuenta para calcular directamente los índices de eficiencia definitivos y la alternativa de los estudios de segunda etapa. El análisis se ha realizado para 87 centros públicos de la Comunidad Autónoma extremeña en el curso 2001-2002.

## 1. INTRODUCCIÓN

En la última década se han llevado a cabo una serie de estudios en nuestro país que han tenido como objetivo analizar la eficiencia de los centros educativos en Educación Secundaria<sup>1</sup> aplicando la técnica envolvente de datos (DEA). Todos estos trabajos inciden en la importancia que tiene la inclusión de las variables no controlables por los centros educativos –representados fundamentalmente por las características del alumno y por su status socioeconómico- en los resultados de dichos análisis.

Un repaso de la literatura especializada nos permite destacar dos posibilidades para poder incluir los inputs no controlables en el análisis DEA<sup>2</sup>. Una primera posibilidad es la consideración de los inputs no controlables en el cálculo de los índices de eficiencia, bien en un solo DEA (análisis pionero de Banker y Morey (1986)) o en varios sucesivos (modelo de tres etapas de Fried y Lovell (1996)). La segunda posibilidad es la realización de “*análisis de segunda etapa*” que consisten en ignorar inicialmente el efecto de los inputs no controlables en la evaluación de los centros, para corregir posteriormente los índices calculados incorporando los efectos de dichas variables.

El objetivo de este trabajo es estudiar la sensibilidad de los índices de eficiencia estimados para los distintos centros educativos a las diferentes alternativas propuestas en la literatura para incluir los factores exógenos en el análisis. Concretamente, nos centraremos en dos cuestiones fundamentales. Contando con una información muy rica suministrada por unas encuestas sobre estas variables consideraremos, por un lado, la posibilidad de incluir directamente las variables más relevantes en el análisis de eficiencia o de utilizar, alternativamente, el análisis de componentes principales para sintetizar la información recogida en dichas variables. Por otro lado, la comparación entre el enfoque que considera que estos factores deben tenerse en cuenta para calcular directamente los índices de eficiencia definitivos y la alternativa de los “estudios de segunda etapa”.

## 2. INFORMACIÓN DISPONIBLE Y VARIABLES UTILIZADAS

Para el análisis que pretendemos realizar contamos con datos referentes a 87 centros públicos de educación secundaria de la Comunidad Autónoma de Extremadura. En concreto, la información disponible hace referencia a las notas obtenidas por los alumnos en la Prueba de

---

<sup>1</sup> La mayor parte de ellos están orientados a tratar de solventar las dificultades relativas al desconocimiento de la función de producción educativa y a la identificación de los factores que influyen sobre los resultados escolares. Véase, por ejemplo, Pedraja y Salinas (1996), Mancebón (1996) o Muñiz (2000). Para una revisión de la literatura de la eficiencia en el ámbito educativo, véase Worthington (2001)

<sup>2</sup> Para una revisión más detallada de las diferentes técnicas puede consultarse Muñiz (2001).

Acceso a la Universidad (PAU) en junio de 2002, así como el número de profesores existentes en cada centro y los gastos de gestión de los mismos.

A partir de los datos de la PAU hemos construido dos variables que reflejan el output educativo de los centros. Por un lado, tenemos la variable denominada “OUTPUT1” que refleja el porcentaje de alumnos aprobados respecto a los matriculados a principios de curso<sup>3</sup> como indicador cuantitativo, y por otro lado, “OUTPUT2”, la nota media de los alumnos que aprobaron la PAU, que recoge una información de carácter cualitativo. Por lo que respecta a los inputs sobre los que tienen el control los centros, construimos dos variables a partir de los datos disponibles: “PROFESORES” que refleja el número total de profesores en el centro por cada 100 alumnos y “GASTOS” que representa el gasto total<sup>4</sup> por alumno realizado por el centro durante el ejercicio 2001.

En relación con los inputs no controlables, contamos con la información procedente de una encuesta realizada por el Servicio de Inspección de la Consejería de Educación de la Junta de Extremadura en todos los centros objeto de la evaluación, la cual consta de 54 preguntas, a partir de las cuales hemos extraído doce variables que vienen dadas por el porcentaje de alumnos que cumplen una serie de condiciones: 1) Alumnos que estudian más de 10 horas semanales; 2) alumnos que llevan todas las asignaturas al día; 3) alumnos que asisten siempre a clase; 4) alumnos con buenas notas en el curso anterior; 5) alumnos que se consideran capacitados para cursar con buenas notas una carrera universitaria; 6) alumnos cuyos padres consideran que están capacitados para terminar con éxito una carrera universitaria; 7) alumnos que no han cambiado de centro; 8) alumnos cuyos padres tienen unos ingresos mensuales superiores a los 1.800 euros; 9-10) alumnos cuyos padres tienen una profesión que requiera una alta cualificación y 11-12) alumnos cuyos padres tienen una carrera universitaria.

Dadas las características del DEA, en el que existe una relación inversa entre el número de variables utilizadas y el poder de discriminación de la técnica, no resulta factible introducir los valores de todos los inputs no controlables con los que contamos. Ante este problema, la literatura plantea dos alternativas. Por un lado, se puede renunciar a una parte de la información con la que contamos e introducir un número reducido de variables en el análisis, y, por otro, se puede comprimir la información en un número reducido de variables mediante el uso del análisis de componentes principales (ACP)<sup>5</sup>.

---

<sup>3</sup> Hemos preferido utilizar este porcentaje en lugar del número de alumnos aprobados respecto a los presentados, para evitar el comportamiento especulativo por parte de los centros a la hora de determinar el número de alumnos que pueden realizar el examen de acceso a la universidad.

<sup>4</sup> Para construir esta variable se ha deducido del gasto total del ejercicio los recursos destinados a edificios y otras construcciones, ya que consideramos que esta partida de gasto varía mucho entre diferentes ejercicios en función de las circunstancias particulares de cada centro.

<sup>5</sup> Smith y Mayston (1987) fueron los primeros que recomendaron el uso de esta técnica para reducir el número de elementos no discrecionales en un contexto de evaluación de la eficiencia de centros. En XII Jornadas de la Asociación de Economía de la Educación

El ACP nos ha permitido identificar cuatro componentes que recogen el 73,2% % del total de la información original. El primer factor estaría compuesto por las cinco variables relativas a los ingresos familiares, los estudios y la profesión de los padres, es decir, por aquéllas que determinan el “*entorno socioeconómico del alumno*”, mientras que los otros tres factores recogen diferentes variables acerca de la opinión del alumno y de sus padres acerca de su éxito futuro (2º componente) y sobre el esfuerzo realizado (3º y 4º componente). Desde el punto de vista teórico, consideramos que las variables que se recogen en estos tres componentes pueden resumirse básicamente en lo que denominamos “*características propias del alumno*”. Basándonos en este argumento y en el hecho, ya comentado anteriormente, de que el DEA pierde poder de discriminación con la introducción de un mayor número de variables, procedemos a refundir estos tres factores en uno solo, mediante el cálculo de una media ponderada de los mismos. Para su cálculo utilizamos como ponderaciones los porcentajes de la varianza explicada por cada factor. De este modo, los dos inputs no controlables que utilizaremos en el análisis de eficiencia serán el entorno socioeconómico del alumno (primer componente obtenido en el ACP, al que llamaremos “INC1”) y las características propias de éste (combinación lineal de los otros tres factores procedentes del ACP, que denominamos “INC2”).

Como alternativa a la opción del análisis de componentes principales, podemos utilizar las propias variables extraídas de la encuesta. Con el fin de que la dimensión del modelo no afecte a los resultados, elegimos dos variables. De acuerdo con la teoría, debería de tratarse de variables no correlacionadas y que recogiesen información acerca de los aspectos principales que afectan a los resultados de los estudiantes, el entorno socioeconómico y sus propias características. De este modo, las dos variables elegidas, siguiendo el criterio de obtener la máxima información no redundante<sup>6</sup>, fueron “INGRESOS”, como indicador de familias con rentas mensuales elevadas, y “BNOTAS”, como aproximación a las propias capacidades de los alumnos.

Como ya mencionamos en el apartado anterior, existen dos alternativas para incluir los factores exógenos en el análisis de eficiencia. La primera, apuesta por el uso de los factores exógenos para construir los índices de eficiencia de los centros. En nuestro caso utilizaremos la metodología de Banker y Morey (1986)<sup>7</sup>.

---

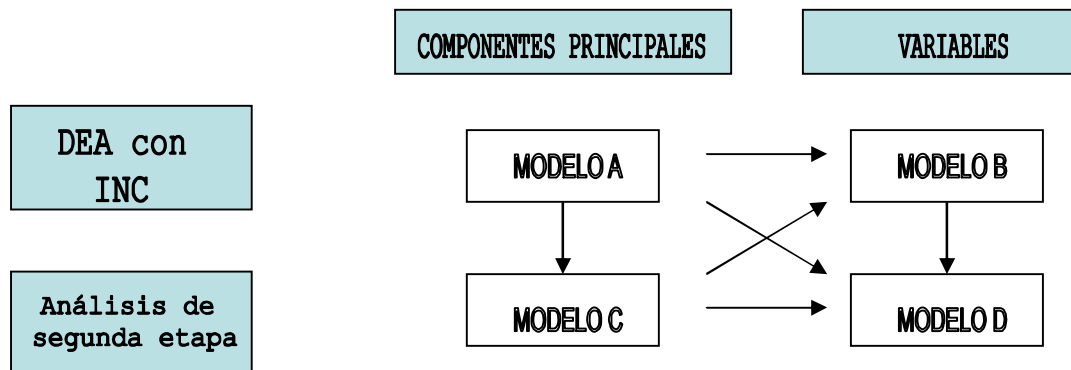
España, Mancebón (1996) utilizó esta técnica en su tesis doctoral sobre la evaluación de la eficiencia de los institutos de educación secundaria de Zaragoza.

<sup>6</sup> Es decir, que no muestran correlación entre ellas y con una influencia significativa sobre el output. Este criterio fue el seguido por Muñiz (2000) para incluir factores exógenos en su tesis doctoral.

<sup>7</sup> El modelo DEA utilizado tiene una orientación de minimización de inputs y utiliza rendimientos de escala constantes.

La segunda consiste en los denominados análisis de segunda etapa<sup>8</sup>. El primer paso consistirá en la realización de un DEA sin tener en cuenta los factores exógenos, para, posteriormente, utilizar los índices de eficiencia obtenidos como variable dependiente en una regresión Tobit en la que se incluyen como variables explicativas los inputs no controlables. Aunque esta opción permite incluir un mayor número de variables explicativas en la regresión, como nuestro objetivo es poder comparar modelos con una estructura similar, hemos preferido utilizar en nuestra regresión las mismas variables introducidas en el análisis DEA (en el primer caso, las dos variables procedentes del análisis de componentes principales y, en el segundo, “ingresos” y “bnotas”)<sup>9</sup>.

Combinando las dos posibilidades descritas anteriormente nos encontramos con cuatro modelos alternativos (A, B, C y D), recogidos en el siguiente esquema:



### 3.- COMPARACIÓN DE RESULTADOS Y CONCLUSIONES

Para comparar los resultados ofrecidos por los cuatro modelos podemos, en primer lugar, examinar los índices de eficiencia estimados por cada uno de ellos. En la tabla 1, se ofrece información sobre el número de unidades eficientes y la eficiencia media obtenida en cada uno de los modelos, pudiendo observarse notables diferencias entre ellos.

**Tabla 1: NÚMERO DE UNIDADES EFICIENTES Y EFICIENCIA MEDIA**

|          | Nº UNIDADES EFICIENTES (%) | EFICIENCIA MEDIA | EFICIENCIA MÍNIMA | EFICIENCIA MÁXIMA |
|----------|----------------------------|------------------|-------------------|-------------------|
| MODELO A | 20 (23%)                   | 87,5%            | 58,3%             | 100%              |
| MODELO B | 17 (19,5%)                 | 84,9%            | 50,7%             | 100%              |
| MODELO C | 11 (12,6%)                 | 79,7%            | 49,6%             | 100%              |
| MODELO D | 16 (18,4%)                 | 84,5%            | 42,9%             | 100%              |

<sup>8</sup> Algunos trabajos que utilizan esta técnica son Ray (1991) y McCarty y Yaisawarng (1993).

<sup>9</sup> Todas las variables resultaron significativas, al 99%, en la explicación de los índices de eficiencia.

En las tablas 2 y 3 se ofrecen los coeficientes de correlación de Pearson y de Spearman de los índices calculados por los distintos modelos.

**Tabla 2: COEFICIENTES DE CORRELACIÓN DE PEARSON**

|          | MODELO A | MODELO B | MODELO C | MODELO D |
|----------|----------|----------|----------|----------|
| MODELO A | 1        |          |          |          |
| MODELO B | 0,739    | 1        |          |          |
| MODELO C | 0,491    | 0,606    | 1        |          |
| MODELO D | 0,425    | 0,547    | 0,918    | 1        |

**Tabla 3: COEFICIENTES DE CORRELACIÓN DE RANGOS DE SPEARMAN**

|          | MODELO A | MODELO B | MODELO C | MODELO D |
|----------|----------|----------|----------|----------|
| MODELO A | 1        |          |          |          |
| MODELO B | 0,723    | 1        |          |          |
| MODELO C | 0,452    | 0,588    | 1        |          |
| MODELO D | 0,429    | 0,542    | 0,944    | 1        |

Puede destacarse que, aunque se siguen produciendo notables diferencias en los índices estimados por cada uno de los modelos alternativos, las menores diferencias se encuentran entre aquellos que utilizan la misma metodología para llevar a cabo la inclusión de los inputs no discrecionales, es decir, entre los dos modelos DEA (A y B) y entre los dos modelos de segunda etapa (C y D).

Una conclusión similar se alcanza si analizamos las divergencias existentes entre los índices de eficiencia calculados por los cuatro modelos. En la tabla 4 se ofrece información sobre el porcentaje de unidades que presentan unos niveles de eficiencia muy similares en los diferentes modelos (comparándolos de dos en dos)<sup>10</sup>.

**PORCENTAJE DE UNIDADES CON DIVERGENCIA ENTRE INDICES < 5 %**

|          | MODELO A | MODELO B | MODELO C | MODELO D |
|----------|----------|----------|----------|----------|
| MODELO A | -        |          |          |          |
| MODELO B | 68 %     | -        |          |          |
| MODELO C | 30 %     | 41 %     | -        |          |
| MODELO D | 34 %     | 40 %     | 52 %     | -        |

Los resultados obtenidos en este trabajo muestran que los índices calculados por el modelo DEA son muy sensibles a la forma en que las variables no controlables son incluidas en el análisis de eficiencia. La opción de incluir directamente las variables exógenas o la de agruparlas en componentes principales afecta de forma considerable a los índices de eficiencia estimados, siendo mayor las discrepancias que se producen en el modelo convencional de Banker y Morey (1986) que en el modelo alternativo de dos etapas. La abundante literatura empírica que se ha realizado, tanto en el ámbito internacional como en nuestro país, en este campo ha optado por incluir las variables no controlables de una u otra forma, sin prestar la debida atención a este problema, por lo que debería profundizarse en su estudio en el futuro.

**4. REFERENCIAS BIBLIOGRÁFICAS**

BANKER, R.D. y MOREY, R.C. (1986): "Efficiency Analysis for exogenously fixed inputs and outputs", *Operations Research*, vol. 34, nº 4, pp. 513-521.

FRIED, H.O. y LOVELL, C.A.K. (1996): "Searching the zeds", ponencia presentada en el II Georgia Productivity Workshop.

GOLANY, B. y ROLL, Y. (1993): "Some Extensions of Techniques to Handle Non-Discretionary Factors in Data Envelopment Analysis", *The Journal of Productivity Analysis* 4, pp. 419-432.

MANCIBÓN, M.J. (1996): "La evaluación de la eficiencia de los centros educativos públicos", Tesis Doctoral, Universidad de Zaragoza.

McCARTY, T. y YAISAWARNG, S (1993): "Technical efficiency in New Jersey School districts", en FRIED, H., LOVELL, C.A.K. y SCHMIDT, S. (ed.): "The Measurement of Productive Efficiency: Techniques and Applications", Oxford University Press, New York.

<sup>10</sup> Como indicador se ha tomado como referencia una medida *ad hoc*, que determina qué porcentaje de unidades tienen una divergencia inferior a un 5 %.

MUÑIZ, M. (2000): “Eficiencia técnica e inputs no controlables. El caso de los institutos asturianos de educación secundaria”, Tesis Doctoral, Universidad de Oviedo.

MUÑIZ, M. (2001): “Introducción de variables de control en modelos DEA”, en ÁLVAREZ, A. (coord.): “La medición de la eficiencia y la productividad”, Editorial Pirámide.

PEDRAJA, F. y SALINAS, J. (1996): “Eficiencia del gasto público en educación secundaria: una aplicación de la técnica envolvente de datos”, *Hacienda Pública Española*, nº 138, pp. 87-95.

RAY, S.C. (1991): “Resource use efficiency in public schools: A study of Connecticut data”, *Management Science*, vol. 37, nº 12, págs. 1.620-1.628.

SMITH, P. y MAYSTON, D. (1987): “Measuring efficiency in the public sector”, *OMEGA International Journal of Management Science*, nº 15, pp. 181-189.

WORTHINGTON, A.C. (2001): “An Empirical Survey of Frontier Efficiency Measurement techniques in Education”, *Education Economics*, vol. 9, nº 3.