

EFICIENCIA Y GESTIÓN EDUCATIVA. UN ANÁLISIS DE LAS DIFERENCIAS ENTRE LAS COMUNIDADES AUTÓNOMAS ESPAÑOLAS EN PISA 2006

José Manuel Cordero Ferrera
Eva Crespo Cebada

Universidad de Extremadura

Daniel Santín González

Universidad Complutense de Madrid

RESUMEN

La reciente publicación del Informe PISA 2006 ha puesto de manifiesto la existencia de diferencias notables en cuanto a las puntuaciones obtenidas por los alumnos en las pruebas de aptitud realizadas entre las regiones españolas que se han sometido a este estudio. En este trabajo se pretende explotar el gran volumen de información que ofrece esta base de datos con el propósito de identificar el posible origen de estas divergencias. Con este fin, se plantea un análisis de eficiencia en distintos niveles (alumnos y escuelas) en el que se incluye información sobre las variables representativas de los principales factores productivos que pueden condicionar los resultados. La técnica empleada para la realización de este análisis es el Análisis Envoltante de Datos (DEA).

Palabras clave: eficiencia, educación, comunidades autónomas, PISA.

1. INTRODUCCIÓN

Desde su aparición en el año 2000, la publicación de los resultados del Informe PISA (*Programme for International Student Assessment*), iniciativa impulsada por la OCDE para evaluar el conocimiento de los alumnos a nivel internacional, ha generado en nuestro país un clima de inquietud y preocupación ante los decepcionantes resultados cosechados, especialmente si se tiene en cuenta que el presupuesto educativo español es superior al de la mayoría de los demás países participantes. Así, es habitual que tras la publicación de estos resultados, asistamos a una intensificación del debate político, normalmente polarizado en el nivel estatal, sobre cuáles deben ser las medidas más apropiadas para poder acortar distancias con el resto de países, aunque en realidad las responsables de las políticas educativas en nuestro país son las Comunidades Autónomas¹.

En este contexto cobra una gran relevancia el hecho de que en la última oleada del Informe PISA, realizada en el año 2006, diez Comunidades Autónomas españolas decidieran participar con una muestra ampliada, lo que permite extraer conclusiones independientes de las obtenidas por el resto de España para cada una de ellas². La tabla 1 muestra las divergencias existentes en cuanto a los resultados obtenidos por los alumnos españoles pertenecientes a cada una de ellas. Sin embargo, una evaluación basada únicamente en los resultados resulta a todas luces un enfoque demasiado simplista, ya que éstos están condicionados en gran medida por los factores educativos utilizados.

En este trabajo se pretende explotar el gran volumen de información que ofrece esta base de datos sobre multitud de aspectos que intervienen en el proceso de aprendizaje mediante un enfoque basado en la evaluación de la eficiencia, con el objetivo de realizar un diagnóstico de las posibles deficiencias en la organización y gestión de los sistemas educativos de ca-

¹ En 2005 el gasto educativo gestionado por el conjunto de las Comunidades Autónomas representaba un 88 % del total, mientras que el Estado y las Corporaciones Locales se reparten el 12 % restante.

² En realidad, en el año 2003 ya hubo tres Comunidades Autónomas que participaron en la evaluación (Castilla y León, Cataluña y País Vasco).

da región. Además, en una segunda etapa se indagará acerca de las fuentes que explican posibles comportamientos ineficientes. Por último, estudiaremos que parte de la ineficiencia no explicada es debida al alumno y a las escuelas dentro en cada Comunidad.

Tabla 1. Resultados en PISA 2006 por CCAA ordenadas según el resultado en Ciencias

	Ciencias	Lectura	Matemáticas
Castilla y León	520	478	515
La Rioja	520	492	526
Aragón	513	483	513
Navarra	511	481	515
Cantabria	509	475	502
Asturias	508	477	497
Galicia	505	479	494
País Vasco	495	487	501
Cataluña	491	477	488
España	488	461	480
Resto de España*	480	450	469
Andalucía	474	445	463

La información para llevar a cabo este estudio se encuentra disponible a nivel de alumno (datos individuales), lo que supone una importante ventaja a la hora de analizar e interpretar los resultados de las estimaciones (Summers y Wolfe, 1977; Hanushek *et al.*, 1996). Sin embargo, dado que la mayor parte de los estudios realizados en este campo han sido realizados utilizando datos agregados a nivel escolar, debido fundamentalmente, a la escasez de bases de datos disponibles a nivel de alumno (Muñiz 2002, Mancebón, 1999 o Cordero *et al.*, 2005), hemos optado por utilizar una doble perspectiva para la realización del análisis: datos a nivel de alumno y a nivel de escuela. A nuestro modo de ver, este doble enfoque permite enriquecer las conclusiones obtenidas, pues permitirá abordar los problemas asociados a la agregación de datos en el contexto educativo.

2. METODOLOGÍA

Siendo el objetivo de este trabajo encontrar la relación existente entre recursos productivos y resultados educativos, el enfoque utilizado se basa en la utilización de la función de producción educativa definida por Levin (1974) y Hanushek (1979) según la cual el output, frecuentemente medido como el resultado en una prueba de conocimientos objetiva, depende de varios factores: las características socioeconómicas en el hogar, los inputs escolares tales como material educativo o la infraestructura del centro, las influencias de los compañeros (efecto *peer group*), la capacidad innata del alumno y su esfuerzo.

Esta función puede ser estimada asumiendo que existen comportamientos ineficientes por parte de las escuelas o de los alumnos, los cuales pueden deberse a motivos tan diversos como la estructura del sistema establecida en la legislación vigente, la forma de organizar los recursos o la propia motivación de los agentes implicados en el proceso, representados fundamentalmente por los alumnos, los padres y los profesores. A pesar de la gran cantidad de estudios que, durante las últimas décadas, se han ocupado de identificar y cuantificar el efecto de éstos y otros aspectos sobre el proceso educativo, los resultados siguen siendo inconclusos.

En el contexto de este trabajo, consideramos tres tipos de variables: los conocimientos obtenidos por los alumnos en distintas materias (outputs), el vector de variables educativas imprescindibles para llevar a cabo la educación, cuyo efecto sobre los resultados ha de ser positivo, esto es, una mayor dotación de cualquiera de estas variables tiene un efecto positivo sobre los resultados y, finalmente, un conjunto de variables que tienen influencia sobre el pro-

ceso, pero sobre las cuales no se conoce *a priori* si dicho efecto puede ser positivo, negativo o inexistente (variables explicativas de la ineficiencia).

El método que utilizamos para medir la eficiencia técnica es el Análisis Envolvente de Datos (DEA), una técnica de carácter no paramétrico, cuya flexibilidad permite adaptarse con gran facilidad a un contexto como el educativo en el que resulta muy difícil modelizar las relaciones existentes entre las variables. Esta técnica ha sido ampliamente utilizada en el contexto de la enseñanza no universitaria desde los trabajos pioneros de Bessent y Bessent (1980), Charnes *et al.* (1981) y Bessent *et al.* (1982), en los que se trabajaba con datos agregados, hasta otros más recientes, en los que se trabaja con datos individuales (Thanassoulis, 1999; Silva y Thanassoulis (2001); Jorge y Santín, 2007).

El modelo DEA nos permite obtener un índice de eficiencia global a partir de la información relativa a los outputs del proceso productivo y los inputs que intervienen en el proceso. Una vez obtenido el índice de eficiencia, en una segunda etapa se puede explicar su valor a partir de un vector de variables $Z = (z_1, z_2, \dots, z_L)$ de determinadas variables individuales, escolares y del sistema educativo que, sin estar directamente relacionadas con la producción educativa, pueden afectar al proceso de aprendizaje: $\theta_i = f(Z_i, \beta_i)$

A partir de los valores de los parámetros estimados $\hat{\beta}_i$, se pueden identificar las variables que influyen en el nivel de ineficiencia e incluso se puede realizar una predicción de cuál debería ser la ineficiencia: $\hat{\theta}_i = f(Z_i, \hat{\beta}_i) + \hat{\varepsilon}_i$

La principal crítica a la que ha tenido que hacer frente este enfoque es el incumplimiento del supuesto de independencia de los errores, puesto que el índice obtenido inicialmente mediante DEA tiene carácter relativo, es decir, incorpora información sobre todas las unidades que componen la muestra (Xue y Harker, 1999; Simar y Wilson, 2007), lo que provoca que el término de error y el vector Z estén correlacionados entre sí. La solución habitual para evitar este tipo de problemas consiste en la utilización de técnicas de bootstrap, mediante las cuales se obtienen estimaciones de los intervalos de confianza y se evitan los problemas de sesgo anteriormente mencionados. En nuestro caso, hemos decidido obviar la utilización de estas técnicas basándonos en que contamos con cinco valores extraídos de la distribución estimada de los resultados en lugar de analizar un único resultado (véase la explicación de los valores plausibles en el apartado posterior), lo que permite trabajar con intervalos de confianza para las estimaciones de la eficiencia desde el principio, haciendo innecesaria la aplicación de métodos de remuestreo.

Una vez corregido el nivel de ineficiencia del alumno existe un término $\hat{\theta}_i - \theta_i = \hat{\varepsilon}_i$ no explicado por el vector de variables Z que sería la ineficiencia observada. Si se trabaja a nivel de alumno, esta ineficiencia puede desglosarse en tres niveles: un componente debido a la propia ineficiencia del alumno, una componente debida a la ineficiencia de la escuela y una componente debida a factores aleatorios como la suerte³. De estos tres factores nos interesa especialmente la ineficiencia debida a la escuela o de gestión del centro, ya que es la parte sobre la cual el sector público debería trabajar en mayor medida para su corrección. La descomposición de estos tres factores la realizaremos a través de un análisis de la varianza del término $\hat{\varepsilon}_i$ en el que asumiremos que las diferencias de ineficiencias entre escuelas (*between*) son debidas a éstas mientras que las diferencias intra-escuela (*within*) son debidas al alumno y a

³ Suponemos que los factores aleatorios que influyen en la ineficiencia no explicada $\hat{\varepsilon}_i$ se distribuyen como una $v \sim N(0; \sigma_v^2)$. Consideraremos esta componente parte de la ineficiencia del alumno ya que ambas, ineficiencia y suerte, quedan inexplicadas por el modelo una vez analizada la ineficiencia debida a la escuela.

factores aleatorios (Perelman y Santín, 2005): $\hat{S}_{\varepsilon_{is}}^2 = \hat{S}_{\varepsilon_s, B}^2 + \hat{S}_{\varepsilon_s, W}^2$

Este análisis permite diferenciar el porcentaje de la ineficiencia no explicada $\hat{S}_{\varepsilon_{is}}^2$ que es debida al alumno y a factores aleatorios de forma separada a la ineficiencia de la escuela. Así, las diferencias en la eficiencia media entre escuelas $\hat{S}_{\varepsilon_s, B}^2$ están asociadas con las características del profesorado, los métodos pedagógicos empleados o la interrelación entre padres y responsables educativos, mientras que las diferencias entre alumnos dentro de cada escuela son atribuibles a la dedicación y esfuerzo de éstos.

3. BASE DE DATOS Y VARIABLES SELECCIONADAS

3.1. Base de datos

La muestra utilizada en esta investigación procede de la tercera oleada del proyecto PISA, realizada en el año 2006, especializada en Ciencias. Concretamente, se utiliza la información relativa a los alumnos españoles que realizaron la prueba, un total de 19.605 pertenecientes a 685 centros educativos, cuya una distribución por Comunidades Autónomas es la que se muestra en la Tabla 2.

Tabla 2. Estructura de la base de datos utilizada

Región	Alumnos	Escuelas	Públicos	Concertados	Privados
Andalucía	1.463	51	37	13	1
Aragón	1.526	51	31	16	4
Asturias	1.579	53	31	14	8
Cantabria	1.496	53	31	19	3
Castilla y León	1.512	52	31	17	4
Cataluña	1.527	51	29	11	10
Galicia	1.573	53	36	11	6
La rioja	1.333	45	22	20	3
Navarra	1.590	52	30	19	3
País Vasco	3.929	150	63	83	4
Resto España	2.077	74	44	20	10
Total	19.605	685	385	243	57

Una de las grandes ventajas que ofrece el proyecto PISA es que no evalúa las capacidades o destrezas mediante una única puntuación, sino que cada alumno recibe una puntuación en cada prueba en una escala continua. Además, dado que la escuela, el hogar y el contexto socioeconómico son variables claves para comprender el mundo educativo, el Proyecto PISA recoge una extensa base de datos en torno a estas variables, obtenidas a partir de dos cuestionarios, uno completado por los propios alumnos y otro por los directores de los centros educativos.

3.2. Variables seleccionadas

Para poder llevar a cabo la evaluación de la eficiencia en el contexto de nuestra evaluación se han seleccionado tres tipos de variables siguiendo los criterios establecidos en la sección previa: outputs, inputs y variables explicativas de la ineficiencia.

Como indicadores del output del proceso educativo, se utilizan los resultados obtenidos por los alumnos en las tres materias evaluadas en PISA: matemáticas, comprensión lectora y ciencias. No obstante, a diferencia de lo que ocurre en la mayor parte de las pruebas estandarizadas de conocimientos, estos resultados no están representados por un único indicador, sino por cinco valores plausibles, entendidos éstos como una representación del rango de habilidades que tiene cada estudiante y que se corresponden con valores aleatorios extraídos de la función de distribución de los resultados estimada a partir de las contestaciones obteni-

das en cada prueba⁴. Para que el cálculo de las medidas de la eficiencia sean correctas, realizaremos el análisis cinco veces, una vez para cada terna de valores plausibles, y posteriormente tomar la media de las estimaciones, ya que si se calculase la media de los valores plausibles y a continuación se realizara el análisis introduciríamos un sesgo en la estimación (OCDE, 2005).

En cuanto a las variables seleccionadas como inputs del proceso educativo, se han considerado tres variables, si bien una de ellas sólo puede incluirse en el análisis realizado a nivel de alumno. En primer lugar, se ha seleccionado la variable *ESCS*, un indicador del estatus socioeconómico familiar, construida por los analistas del proyecto PISA agrupando las respuestas a distintas preguntas planteadas en el cuestionario de la evaluación (el nivel educativo más alto de cualquiera de los padres, el índice más alto de ocupación laboral de cualquiera de los padres y un índice de posesiones educativas relacionadas con la economía del hogar). En segundo lugar, se ha incluido un indicador representativo de la calidad de los recursos educativos del centro, *SCMATEDU*, construido a partir de las respuestas del director de la escuela a siete preguntas relacionadas con la disponibilidad de ordenadores para usos didácticos, software educativo, calculadoras, libros, recursos audiovisuales y equipo de laboratorio. Por último, en el análisis a nivel de alumno se ha incluido una variable representativa del efecto compañeros o *peer-group* (*EFCO*), definida como el nivel medio en la variable *ESCS* de los compañeros que comparten la misma escuela que el alumno evaluado, cuya base teórica reside en el hecho de que el nivel de conocimientos que un alumno puede asimilar depende directamente de las características de sus compañeros en el aula⁵. Dado estas tres variables presentaban valores positivos y negativos, todas ellas fueron reescaladas para que presentaran valores positivos y que de esta forma el modelo DEA pudiera ser resuelto. Asimismo de forma previa a la realización del análisis DEA se comprobó la existencia de una correlación positiva y significativa entre todos los inputs y los outputs introducidos, lo que garantiza una correcta especificación del modelo.

Además de los inputs anteriores, hemos considerado que determinadas características del alumno, la escuela o el sistema educativo pueden afectar al nivel de eficiencia. En particular hemos analizado la posible influencia de las siguientes variables: la titularidad de la escuela, pues consideramos que puede resultar interesante analizar si el carácter público, privado o concertado de la escuela influye en la eficiencia, el tamaño de la escuela (nº total de estudiantes en el centro), el tamaño del aula (el total de alumnos dividido por el total de profesores⁶), el curso académico al que pertenece el estudiante, definido a través de una variable dicotómica que indica si el alumno ha repetido curso o no, el hecho de ser inmigrante, definido a través de dos variables dicotómicas, según sea inmigrante el propio alumno (*inmig1*) o alguno de sus padres (*inmig2*) y, finalmente, la pertenencia a una Comunidad Autónoma, bajo la hipótesis de que los alumnos de determinadas Comunidades pueden ser más eficientes que otros⁷.

4. RESULTADOS

En las tablas 3 y 4 se muestran los valores de la eficiencia técnica medios obtenidos tras resolver los cinco modelos DEA, utilizando una orientación al output y rendimientos variables de escala (Banker, Charnes y Cooper, 1984), tomando los inputs y los outputs en valores plausibles de manera agregada (tomando a la escuela como unidad de análisis) y desagregada.

⁴ Para una revisión de la literatura de los valores plausibles puede acudir a Mislevy (1991) y Mislevy *et al.* (1992). En España, Calero y Escardibul (2007) utilizan los valores plausibles contenidos en PISA 2003 para llevar a cabo una estimación de la función educativa mediante análisis multinivel.

⁵ Para una revisión del efecto de estas variables sobre los resultados véase Betts (2000) o Hanushek *et al.* (2001).

⁶ Los profesores a tiempo parcial computan sólo como medio profesor.

⁷ Se han construido diez variables dicotómicas, una para cada Comunidad Autónoma. Por tanto, la referencia con la que se comparan las regiones es la muestra perteneciente al resto de regiones.

gada (utilizando la información a nivel de alumno e incluyendo el efecto compañeros)⁸.

Una primera aproximación a estos resultados permite identificar a La Rioja, Aragón, Castilla y León, Navarra y Galicia como las más eficientes, mientras que Andalucía, Cataluña y la muestra formada por el resto de regiones se sitúa en el extremo contrario. Asimismo, la comparación entre los valores de ambas tablas, permite apreciar las claras diferencias existentes al utilizar diferentes unidades de análisis, cuya explicación se encuentra principalmente en los grados de libertad considerados en cada modelo. Más interesante resulta el hecho de que las diferencias entre regiones se estrechan cuando se consideran datos individuales. Así, se puede apreciar que las diferencias entre la Comunidad con mejores resultados en ambos casos (La Rioja) y la muestra relativa al resto de regiones, con peores resultados relativos, pasa de ser de aproximadamente el 6 % con datos agregados a una cifra inferior al 4 % con datos individuales. Asimismo, la clasificación de las Comunidades también varía dependiendo de cuál sea la unidad de análisis considerada. Finalmente, nos ha llamado la atención el hecho de que los índices obtenidos con los datos a nivel de alumno apenas se ven influidos por la inclusión de la variable *EFCO*⁹, lo que nos lleva a pensar que la varianza intra-escuela de la variable ESCS

Tabla 3. Resultados del análisis de eficiencia realizado a nivel de centros escolares

CC.AA.	Centros	Vp1	Vp2	Vp3	Vp4	Vp5	Media
La rioja	45	92,52	92,60	92,37	92,26	92,04	92,36
Castilla y león	52	92,15	92,22	91,38	91,64	91,36	91,75
Aragón	51	90,74	90,99	90,79	90,65	90,35	90,70
Navarra	52	90,66	91,11	90,49	90,61	90,41	90,66
Galicia	53	90,41	90,86	89,76	90,59	90,05	90,33
Asturias	53	90,14	90,22	89,89	89,94	89,46	89,93
Cantabria	53	90,19	90,25	89,38	89,78	89,28	89,78
País vasco	150	89,41	89,65	88,90	88,83	88,97	89,15
Cataluña	51	87,45	87,95	87,48	87,37	87,11	87,47
Andalucía	51	87,52	88,06	86,61	87,55	87,13	87,37
Otras	74	86,67	87,06	86,21	86,22	86,26	86,48
Total	685	89,63	89,91	89,19	89,35	89,14	89,44

Tabla 4. Resultados del análisis de eficiencia realizado a nivel de alumnos

CC.AA.	ALUMNOS	VP1	VP2	VP3	VP4	VP5	Media
La Rioja	1.333	69,79	69,77	72,10	70,70	67,44	69,96
Aragón	1.526	68,74	68,95	71,36	69,81	66,63	69,10
Galicia	1.573	68,91	69,03	70,94	69,79	66,61	69,06
Castilla y León	1.512	68,86	68,82	70,87	69,51	66,40	68,89
Navarra	1.590	68,39	68,75	70,77	69,43	66,24	68,72
Cantabria	1.496	68,02	68,06	70,01	68,97	65,58	68,13
País Vasco	3.929	67,75	67,91	70,15	68,70	65,50	68,00
Asturias	1.579	67,84	67,64	70,19	68,74	65,40	67,96
Cataluña	1.527	66,67	66,98	69,39	67,84	64,61	67,10
Andalucía	1.463	66,65	66,64	68,17	67,18	64,29	66,59
Otras	2.077	65,94	66,02	67,96	66,64	63,68	66,05
Total	19.605	66,65	66,64	68,17	67,18	64,29	66,59

⁸ El valor de la eficiencia técnica con el valor plausible 1, sería el obtenido a partir de los tres outputs (notas de matemáticas, lectura y ciencias) asociados a los valores plausibles 1 y sus inputs (que serían los mismos para cada estimación), y así sucesivamente.

⁹ La correlación entre los índices calculados con y sin la variable EFCO es de 0,98.

Tabla 5. Resultados de las regresiones de segunda etapa con datos agregados

Variables	V. PLAUSIBLES 1			V. PLAUSIBLES 2			V. PLAUSIBLES 3			V. PLAUSIBLES 4			V. PLAUSIBLES 5			MEDIA V. PLAUSIBLES		
	Coef.	t	Prob.	Coef.	t	Prob.	Coef.	t	Prob.	Coef.	t	Prob.	Coef.	t	Prob.	Coef.	t	Prob.
Constante	79,043	68,710	0,000	79,599	68,587	0,000	78,272	65,795	0,000	78,589	68,135	0,000	79,065	68,628	0,000	78,914	67,971	0,000
Tipo Centro	-0,006	-0,012	0,990	0,019	0,040	0,968	0,259	0,546	0,586	0,116	0,251	0,802	0,029	0,062	0,951	0,083	0,177	0,859
Tamaño Centro	0,001	2,292	0,022	0,002	2,635	0,009	0,002	2,846	0,005	0,002	2,856	0,004	0,001	2,441	0,015	0,002	2,614	0,011
Tamaño Clase	-0,001	-0,012	0,991	-0,015	-0,205	0,838	-0,032	-0,415	0,678	-0,019	-0,252	0,801	-0,019	-0,250	0,803	-0,017	-0,227	0,822
Curso	11,052	7,751	0,000	10,946	7,610	0,000	11,240	7,624	0,000	10,749	7,519	0,000	10,717	7,506	0,000	10,941	7,602	0,000
Inmig1	-6,163	-1,239	0,216	-5,535	-1,103	0,270	-4,121	-0,801	0,423	-5,546	-1,112	0,266	-2,221	-0,446	0,656	-4,717	-0,940	0,366
Inmig2	1,672	0,387	0,699	0,397	0,091	0,927	-0,575	-0,129	0,898	1,160	0,268	0,789	-2,092	-0,484	0,629	0,112	0,027	0,788
Andalucía	1,205	1,257	0,209	1,319	1,363	0,173	0,833	0,839	0,402	1,736	1,805	0,072	1,211	1,260	0,208	1,261	1,305	0,213
Aragón	3,912	4,191	0,000	3,730	3,962	0,000	4,385	4,544	0,000	4,283	4,577	0,000	3,865	4,135	0,000	4,035	4,282	0,000
Asturias	2,755	2,937	0,003	2,417	2,553	0,011	2,932	3,022	0,003	3,039	3,231	0,001	2,495	2,655	0,008	2,728	2,880	0,005
Cantabria	3,615	3,899	0,000	3,266	3,492	0,001	3,262	3,402	0,001	3,684	3,963	0,000	3,068	3,304	0,001	3,379	3,612	0,000
Cast. Y León	5,081	5,425	0,000	4,697	4,971	0,000	4,718	4,871	0,000	5,037	5,363	0,000	4,626	4,931	0,000	4,832	5,112	0,000
Cataluña	-0,032	-0,034	0,973	0,089	0,094	0,925	0,420	0,431	0,667	0,378	0,399	0,690	0,042	0,045	0,964	0,179	0,187	0,844
Galicia	4,050	4,301	0,000	4,080	4,295	0,000	3,864	3,968	0,000	4,717	4,996	0,000	4,037	4,281	0,000	4,150	4,368	0,000
La Rioja	5,758	5,918	0,000	5,461	5,564	0,000	6,078	6,040	0,000	6,005	6,155	0,000	5,649	5,797	0,000	5,790	5,895	0,000
Navarra	2,917	3,060	0,002	2,931	3,048	0,002	3,104	3,149	0,002	3,324	3,478	0,001	3,006	3,149	0,002	3,056	3,177	0,002
Pais Vasco	1,350	1,774	0,077	1,141	1,486	0,138	1,168	1,484	0,138	1,223	1,602	0,110	1,261	1,655	0,098	1,229	1,600	0,112
Nº Observaciones	685			685			685			685			685			685		

Tabla 6. Resultados de las regresiones de segunda etapa con datos desagregados

Variables	V. PLAUSIBLES 1			V. PLAUSIBLES 2			V. PLAUSIBLES 3			V. PLAUSIBLES 4			V. PLAUSIBLES 5			MEDIA V. PLAUSIBLES		
	Coef.	t	Prob.	Coef.	t	Prob.	Coef.	t	Prob.	Coef.	t	Prob.	Coef.	t	Prob.	Coef.	t	Prob.
Constante	59,810	207,262	0,000	60,420	206,528	0,000	60,280	206,980	0,000	59,763	206,447	0,000	58,484	204,960	0,000	59,751	206,435	0,000
Tipo Centro	-0,738	-4,857	0,000	-0,760	-4,930	0,000	-0,293	-1,912	0,056	-0,609	-3,997	0,000	-0,876	-5,831	0,000	-0,655	-4,305	0,011
Tamaño Centro	0,001	4,028	0,000	0,001	3,236	0,001	0,001	7,366	0,000	0,001	5,245	0,000	0,000	2,539	0,011	0,001	4,483	0,002
Tamaño Clase	0,044	1,791	0,073	0,031	1,236	0,216	0,032	1,314	0,189	0,051	2,076	0,038	0,043	1,766	0,077	0,040	1,637	0,119
Curso	9,791	69,010	0,000	9,478	65,896	0,000	10,644	74,333	0,000	10,249	72,005	0,000	9,019	64,288	0,000	9,836	69,106	0,000
Inmig1	-1,612	-4,142	0,000	-1,647	-4,174	0,000	-1,745	-4,442	0,000	-1,593	-4,080	0,000	-1,278	-3,320	0,001	-1,575	-4,032	0,000
Inmig2	-0,727	-2,136	0,033	-0,746	-2,162	0,031	-1,015	-2,955	0,003	-0,936	-2,741	0,006	-0,809	-2,402	0,016	-0,847	-2,479	0,018
Andalucía	1,009	3,286	0,001	0,883	2,836	0,005	0,707	2,283	0,022	0,906	2,941	0,003	0,816	2,688	0,007	0,864	2,807	0,008
Aragón	2,694	8,998	0,000	2,801	9,225	0,000	3,351	11,088	0,000	3,088	10,280	0,000	2,823	9,533	0,000	2,951	9,825	0,000
Asturias	1,482	4,956	0,000	1,169	3,857	0,000	1,817	6,019	0,000	1,693	5,644	0,000	1,304	4,410	0,000	1,493	4,977	0,000
Cantabria	2,145	7,081	0,000	2,053	6,685	0,000	2,237	7,318	0,000	2,449	8,058	0,000	1,910	6,378	0,000	2,159	7,104	0,000
Castilla y León	2,683	8,925	0,000	2,535	8,317	0,000	2,696	8,885	0,000	2,652	8,792	0,000	2,482	8,347	0,000	2,610	8,653	0,000
Cataluña	0,189	0,626	0,531	0,416	1,360	0,174	0,850	2,791	0,005	0,643	2,124	0,034	0,424	1,420	0,156	0,504	1,664	0,180
Galicia	3,330	11,057	0,000	3,275	10,726	0,000	3,537	11,636	0,000	3,606	11,935	0,000	3,182	10,684	0,000	3,386	11,208	0,000
La Rioja	3,585	11,491	0,000	3,463	10,947	0,000	3,945	12,529	0,000	3,809	12,169	0,000	3,473	11,255	0,000	3,655	11,678	0,000
Navarra	1,587	5,293	0,000	1,842	6,061	0,000	1,911	6,318	0,000	1,918	6,377	0,000	1,734	5,851	0,000	1,798	5,980	0,000
País Vasco	0,477	1,946	0,052	0,571	2,298	0,022	0,668	2,702	0,007	0,640	2,601	0,009	0,597	2,465	0,014	0,591	2,402	0,021
Nº Observaciones	19.605			19.605			19.605			19.605			19.605			19.605		

es muy pequeña, lo que sería equivalente a decir que el reparto de los alumnos entre escuelas refleja de manera muy fiel el entorno en el que residen los alumnos, algo que resulta bastante normal con un proceso de asignación escolar basado principalmente en la proximidad al centro de estudios.

Una vez realizado el análisis DEA procedemos a explicar los índices de eficiencia obtenidos por los factores explicativos de la eficiencia. Los resultados obtenidos en las diferentes regresiones estimadas se muestran en las Tablas 5 y 6.

A partir de los valores medios de los cinco análisis Tobit realizados a nivel de escuela y de alumno se pueden extraer un número de conclusiones muy interesantes¹. En primer lugar, se puede apreciar que hay variables que son consideradas explicativas cuando se utilizan datos a nivel de alumno que no lo son cuando la unidad de análisis es la escuela (el tipo de centro o la condición de inmigrante), lo que nos alerta acerca de los problemas que conlleva la agregación de datos a la hora de encontrar variables que ayuden a explicar posibles ineficiencias en los resultados condicionantes de la eficiencia (Santín, 2006). En segundo lugar, los resultados de las regresiones realizadas en ambos niveles coinciden en descartar el tamaño de la clase como una variable explicativa de la ineficiencia. En tercer lugar, existe una clara correspondencia entre el curso en el que se encuentra el alumno y un comportamiento más eficiente, lo cual plantea interrogantes en torno a la conveniencia o no de la política de repetición de curso y de los factores que pueden condicionarla. Por último, el nivel de eficiencia alcanzado por los alumnos y los centros de todas las Comunidades Autónomas es mayor que el de la muestra relativa al resto de regiones, que actúa como referencia en la evaluación, a excepción de Cataluña y País Vasco, las únicas con valores que no son estadísticamente significativos al 99 % a nivel agregado y desagregado.

Una vez analizados los resultados del análisis de eficiencia inicial y el análisis de segunda etapa, cabe preguntarse cuál es el porcentaje de la ineficiencia del alumno que es directamente atribuible a la escuela una vez descontado el efecto de las variables anteriores. Para ello, se ha realizado un análisis de la varianza de los resultados obtenidos a nivel de alumno, a partir del cual se pueden identificar las diferencias entre las eficiencias medias de los alumnos que pertenecen a distintas escuelas (varianza inter-escuelas), atribuibles a la ineficiencia de la gestión de la propia escuela y la varianza entre los alumnos de la misma escuela (varianza intra-escuela).

Los resultados obtenidos en cada región, recogidos en la Tabla 7, ponen de manifiesto que la mayor parte de la ineficiencia detectada depende fundamentalmente del alumno, puesto que la ineficiencia atribuible a la escuela en ningún caso supera el 20 %. Ello denota que la calidad de las escuelas es bastante uniforme en todo el territorio nacional, lo cual simplifica a los padres el problema de la elección escolar. No obstante, se aprecian notables diferencias entre regiones como Andalucía o La Rioja, curiosamente las situadas en extremos opuestos en términos de eficiencia, con valores próximos al 20 % y otras Comunidades como Aragón o Cantabria con valores cercanos al 10 %.

¹ Dadas las mínimas diferencias existentes entre los índices de eficiencia calculados incluyendo y excluyendo a la variable efecto compañeros, hemos utilizado como variable dependiente en las regresiones los índices obtenidos dejando al margen esta variable para mantener un criterio de homogeneidad con las estimaciones realizadas utilizando datos agregados.

Tabla 7. Análisis de la varianza explicada por los alumno y la escuela

CC.AA.	Between	Within	N° Observaciones		F test
	(escuela)	(alumno)	Escuelas	Alumnos	
Andalucía	19,33	80,67	51	1.463	6,840*
Aragón	10,54	89,46	51	1.526	3,576*
Asturias	15,22	84,78	53	1.579	5,218*
Cantabria	10,87	89,13	53	1.496	3,198*
Castilla y León	13,20	86,80	52	1.512	4,215*
Cataluña	11,78	88,22	51	1.527	3,889*
Galicia	12,97	87,03	53	1.573	4,321*
La Rioja	18,95	81,05	45	1.333	6,576*
Navarra	17,13	82,87	52	1.590	6,279*
País Vasco	17,64	82,36	150	3.929	5,748*
Resto España	14,11	85,89	74	2.077	4,736*
Media	14,70	85,30	685	19.605	

5. CONCLUSIONES

El objetivo del trabajo realizado ha consistido en analizar las diferencias existentes en términos de eficiencia en el servicio educativo entre las Comunidades Autónomas españolas, tomando como referencia la información relativa a la última oleada del Informe PISA, realizado en 2006.

La realización del estudio considerando dos posibles unidades de análisis nos ha permitido apreciar que la elección de una de estas dos alternativas condiciona en gran medida los resultados, tanto en lo que se refiere a los valores de los índices de eficiencia obtenidos como a la hora de identificar posibles variables explicativas de la ineficiencia.

A pesar de estas diferencias, los resultados obtenidos para ambos niveles, coinciden en señalar que La Rioja no es sólo la Comunidad con los mejores resultados en PISA 2006, sino que además es la Comunidad que presenta unos mayores valores niveles de eficiencia media. En extremo contrario se sitúa Andalucía, con los peores resultados académicos y los mayores niveles de ineficiencia. No obstante, la inclusión en el análisis de los factores productivos que intervienen en el proceso educativo hace que las diferencias entre regiones se estrechen notablemente, especialmente cuando se utilizan datos desagregados a nivel de alumnos.

6. REFERENCIAS BIBLIOGRÁFICAS

- BANKER, R. D.; CHARNES, A. y COOPER, W. W. (1984). "Models for estimating technical and scale efficiencies in data envelopment analysis", *Management Science* 30 (9), pp. 1078-92.
- BESSENT, A., BESSENT, W., KENNINGTON, J. y REAGAN, B. (1982). "An application of mathematical programming to assess productivity in the Houston independent school district", *Management Science*, 28, pp. 1355-1367.
- BETTS, J. R. y SHKOLNIK, J. L. (2000). "The effects of ability grouping on student achievement and resource allocation in secondary schools", *Economics of Education Review*, 19, pp. 1-15.
- CALERO, J. y ESCARDIBUL, J.O. (2007): "Evaluación de servicios educativos: el rendimiento en los centros públicos y privados medido en PISA-2003", *Hacienda Pública Española*, nº 183 (4/2007), pp. 33-66.
- CORDERO, J. M., PEDRAJA, F. y SALINAS, J. (2005). "Eficiencia en educación secundaria e inputs no controlables: sensibilidad de los resultados ante modelos alternativos". *Hacienda Pública Española / Revista de Economía Pública*, 173 (2), pp. 61-83.
- CHARNES, A., COOPER, W. W. y RHODES, E. (1981). "Evaluating program and managerial efficiency: an application of data envelopment analysis to Program Follow Through", *Management Science*, 27, 668-697.
- HANUSHEK, E. A. (1979). "Conceptual and empirical issues in the estimation of educational production functions." *Journal of Human Resources*, 14, 351-388.

- HANUSHEK, E. A., RIVKIN, S. G. y TAYLOR, L. L. (1996). "Aggregation and the estimated effects of school resources". *The Review of Economics and Statistics*. November 1996. 78 (4), 611-627.
- HANUSHEK, E. A., KAIN, J. F., MARKMAN, J. M. y RIVKIN, S.G. (2001): "Does peer ability affect student achievement?" *Working Paper 8502*, National Bureau of Economic Research.
- JORGE, J. y SANTÍN, D. (2007): "La medición de la eficiencia educativa de los alumnos de 15 años en la Unión Europea", ponencia presentada en el III Congreso de Eficiencia y Productividad (EFIUCO), Córdoba.
- LEVIN, H. M. (1974). "Measuring efficiency in educational production." *Public Finance Quarterly*, 2, 3-24.
- OECD (2005): *PISA 2003 Data Analysis Manual. SPSS users*. Organisation for Economic Cooperation and Development, Paris.
- PERELMAN, S. y SANTÍN, D. (2005). Measuring educational efficiency at student level with parametric stochastic distance functions: an application to Spanish PISA results. Cahiers du CREPP 2005/04. Université de Liège. http://www.ulg.ac.be/crepp/2005_04.pdf
- SANTÍN, D. (2006). "La medición de la eficiencia de las escuelas: una revisión crítica". *Hacienda Pública Española / Revista de Economía Pública* 177 (2), 57-83.
- SILVA y THANASSOULIS (2001). Descomposing school and school type efficiency. *European Journal of Operational Research*, 132, 357-373.
- SIMAR, L. y WILSON, P.W. (2007). "Estimation and Inference in Two-Stage, Semiparametric Models of Production Processes", *Journal of Econometrics*, 136, pp. 31-64.
- SUMMERS, A. A. y WOLFE, B. L. (1977). "Do schools make a difference?" *American Economic Review* 67 (4), 639-652.
- THANASSOULIS, E. (1999). "Setting achievements targets for school children", *Education Economics*, 7(2), 101-119.
- XUE, M. y HARKER, P.T. (1999): "Overcoming the Inherent Dependency of DEA Efficiency Scores: A Bootstrap Approach", Working Paper, Wharton Financial Institutions Center, University of Pennsylvania.