

AURELIA MODREGO RICO (*)
M.ª JESUS SAN SEGUNDO GOMEZ DE C. (*)

1. INTRODUCCION

En este trabajo estimamos ecuaciones de rendimiento escolar utilizando datos referentes a alumnos de Enseñanzas Medias. Al estimar estas ecuaciones intentamos identificar las características de los estudiantes, sus familias, los centros educativos a los que asisten y los estudios que realizan, que muestran una relación estadística significativa con el rendimiento escolar en diversas materias.

Utilizamos estas estimaciones para evaluar la Reforma de las EE.MM. al comparar los resultados obtenidos por sus alumnos con los resultados obtenidos por los alumnos de BUP y FP.

En la próxima sección especificamos las ecuaciones de rendimiento que vamos a estimar. A continuación describimos nuestro conjunto de datos y las variables que se utilizan en este estudio. En la sección 4 presentamos estimaciones de ecuaciones de rendimiento en niveles y en diferencias. En la sección 5 comparamos los estudiantes experimentales únicamente con el grupo de control interno, y en la sección 6 resumimos nuestras conclusiones.

En el Apéndice A se consideran problemas de especificación y en el Apéndice E se describe la construcción de los índices socioeconómico y de habilidad que se han utilizado en este trabajo.

2. ESPECIFICACION DE ECUACIONES DE RENDIMIENTO

La mayor parte de los estudios empíricos de rendimiento escolar (revisados en Modrego y San Segundo) considera ecuaciones de rendimiento de la siguiente forma:

(*) Universidad del País Vasco.

$$(1) \quad R_{ji_t} - R_{ji_{t-1}} = X_{ji} \beta + Z_i \varphi E_{ji} \delta + U_{ji}$$

Es decir, ecuaciones que relacionan la evolución en el rendimiento escolar del período $t - 1$ al t para el estudiante j en el centro i , con las características del estudiante y su familia (X_{ji}), con las características de la escuela (Z_i) a la que asiste y con un término aleatorio (U_{ji}). En casos como el que nos ocupa se añade una variable (E_{ji}) que indica si el individuo pertenece al grupo experimental o no.

Suponemos que esta ecuación de aprendizaje corresponde al siguiente modelo de rendimiento escolar:

$$(2) \quad R_{ji_t} = X_{ji} b_t + Z_i r_t + E_{ji} d_t + \alpha_{ji} + \varnothing_i + V_{ji_t}$$

Estamos, por lo tanto, suponiendo que el rendimiento escolar en un momento dado del tiempo es una función de las características del estudiante y su familia (X_{ji}), de las características de su escuela (Z_i), de su participación, o no, en el experimento (E_{ji}), de las características individuales no observadas (α_{ji}), de las características no observadas de los centros (\varnothing_i) y de un término aleatorio V_{ji_t} , que suponemos no correlacionado con α , \varnothing , X y Z .

En este caso reinterpretamos la ecuación (1) como:

$$(3) \quad R_{ji_t} - R_{ji_{t-1}} = X_{ji} (b_t - b_{t-1}) + Z_i (r_t - r_{t-1}) + E_{ji} (d_t - d_{t-1}) + V_{ji_t} - V_{ji_{t-1}}$$

donde $\beta = b_t - b_{t-1}$, $\varphi = r_t - r_{t-1}$, $\delta = d_t - d_{t-1}$ y $U_{ji} = V_{ji_t} - V_{ji_{t-1}}$.

Seguindo a Chamberlain (1982), interpretamos que los parámetros cambiantes en el tiempo (b_t , r_t , d_t) significan que conforme maduran los estudiantes, sus características familiares y escolares producen diferentes efectos sobre su rendimiento escolar. Por el contrario, suponemos que las variables explicativas permanecen invariantes entre $t - 1$ y t , ya que éste parece ser el caso relevante para nuestro trabajo.

Si estimamos ecuaciones en diferencias —como (3)—, sólo obtendremos coeficientes diferentes de cero para aquellas variables cuya relación con el rendimiento escolar sea diferente en t y en $t - 1$. Si estimamos ecuaciones en niveles —como (2)—, podemos recuperar todos los parámetros (b_t , b_{t-1} , r_t , ...), pero la presencia de características no observadas (α y \varnothing) puede hacer que nuestras estimaciones por MCO de estos parámetros no sean eficientes (si α y \varnothing son factores aleatorios) e incluso que no sean consistentes (si α y \varnothing son factores fijos). En la literatura se considera probable que las ecuaciones (2) incluyan factores no observados α y \varnothing ya que no disponemos de información sobre la inteligencia, la motivación y las aspiraciones de los estudiantes, ni sobre las características de los profesores, sus métodos de trabajo, etc. Esperamos que todas estas variables tengan alguna relación con el

rendimiento escolar de los estudiantes y, por tanto, que formen parte del término de error en (2).

En las ecuaciones en diferencias (3) no se presentan estos problemas si las características no observadas (α y \emptyset) tienen una relación invariante en el tiempo con el rendimiento escolar de los estudiantes. Eliminar los problemas causados por α y \emptyset a través de estimaciones en diferencias es una de las aplicaciones fundamentales de las bases de datos longitudinales que van proliferando en diversas áreas. En este capítulo vamos a presentar estimaciones de ecuaciones de rendimiento en niveles (2) y en diferencias (3). Al especificar estas ecuaciones vamos a intentar utilizar las enseñanzas que se derivan de estudios previos y vamos a aprovechar las posibilidades que ofrece nuestro conjunto de datos.

3. LA MUESTRA: PROBLEMAS DE NO-RESPUESTA Y ELECCION DE VARIABLES

Utilizamos información longitudinal acerca de 7.456 estudiantes que cursaban 1.º de BUP, FP y EXPE en octubre de 1984 en 256 centros escolares públicos y privados. Los datos provienen de las muestras recogidas en 1984 y 1985 para el proyecto de Evaluación de la Reforma de las Enseñanzas Medias del CIDE.

3.a. *Valores perdidos (missing values)*

Nuestra muestra incluye sólo aquellos individuos que participaron en las dos aplicaciones y de los que disponíamos de información completa en cuanto a sus resultados académicos, historial, sexo, edad y población de residencia; incluimos las observaciones que no proporcionan información completa respecto a sus características socioeconómicas y este hecho es tenido en cuenta en la construcción del índice socioeconómico (véase Apéndice E).

Si suponemos que las no-respuestas se distribuyen aleatoriamente entre la población considerada, entonces la reducción en el tamaño de la muestra disminuirá la eficiencia de nuestras estimaciones, pero no introducirá sesgos.

En la Tabla 1 presentamos las estadísticas descriptivas para las variables que empleamos en nuestro estudio, en la muestra original y en la que utilizamos en este trabajo.

3.b. *Elecciones de variables*

VARIABLES explicativas

A) Historial académico de los estudiantes y situación actual

TABLA 1
Estadísticas descriptivas

VARIABLES TESTS	N = 14.065			N = 7.456		
	MEDIA	DESV.S	n.º obs. válidas	MEDIA	DESV.S	n.º obs. válidas
MAT. APL. MAT. APL ₂	4.349 4.984	5.15 3.27	13.658 10.144	4.776 5.117	2.81 3.27	7.456 7.456
MAT. CALC. MAT. CALC ₂	14.269 17.719	6.35 5.80	13.658 9.854	15.497 18.062	6.01 5.67	7.456 7.456
COMLEC COMLEC ₂	12.947 15.287	4.12 3.83	13.536 9.892	13.501 15.459	3.96 3.74	7.456 7.456
ORTOG ORTOG ₂	34.106 35.836	4.41 7.67	13.844 10.142	35.156 36.200	6.85 7.42	7.456 7.456
APTES APTES ₂	39.699 48.840	12.01 12.78	13.785 10.077	41.465 50.488	11.81 12.59	7.456 7.456
APESP APESP ₂	5.712 7.897	5.00 5.27	13.743 10.144	5.922 7.941	4.82 5.21	7.456 7.456
DATMR DATMR ₂	38.215	8.36	13.754	38.745 43.633	(8.02) (7.84)	7.456 7.456
PREESC.	0.850	0.35	13.906	0.870	0.33	7.456
REPET. CURS.	0.410	0.49	13.979	0.326	0.47	7.456
TRABAJA	0.406	0.48	13.923	0.383	0.48	7.456
MENOS 10.000	0.183	0.38	14.063	0.188	0.39	7.456
MADR.BARC	0.111	0.31	14.063	0.102	0.30	7.456
HOMBRE	0.597	0.49	13.990	0.563	0.49	7.456
PUBLICO	0.646	0.47	14.005	0.617	0.48	7.456
BUP EXPER	0.418 0.251	0.49 0.43	10.113 10.113	0.436 0.249	0.50 0.43	7.456 7.456
INDICE NIVELI 1 SOCIOECC				0.000 -0.002 41.500	0.57 0.36 5.59	7.456 7.456 7.456

Incluimos variables ficticias (dummies) que indican si el estudiante asistió a preescolar (PREESC = 1) y si ha repetido alguna vez un curso (REPT. CURS. = 1). En nuestra muestra el 87 por 100 de los estudiantes asistió al colegio antes de comenzar EGB y el 33 por 100 ha repetido curso alguna vez.

Con respecto a la 1.^a variable, estudios del rendimiento escolar en EE.UU. encuentran efectos positivos, pero decrecientes en el tiempo, de haber asistido a educación preescolar para los estudiantes de enseñanza básica (ej.: Henderson-Miezkowski-Sauvageau, 1978) y ningún efecto para los estudiantes de bachiller (San Segundo, 1985). En cuanto a la variable REPT. CURS. se observa que muestra una relación significativa duradera con el bajo rendimiento escolar (San Segundo, 1985).

La situación actual del estudiante viene recogida por las variables que indican si estudia BUP (BUP = 1) y por la «dummy» que indica si el estudiante trabaja o no (TRABAJA = 1). El 43,5 por 100 de los estudiantes de la muestra cursa estudios de Bachiller, el 25 por 100 estudia enseñanzas experimentales y el 31,5 por 100 estudia Formación Profesional. Estamos especialmente interesados en averiguar si el rendimiento de los estudiantes es significativamente diferente en los distintos programas de estudios, una vez que controlamos las diferencias en características de los individuos.

También incluimos dos variables que representan habilidades de los estudiantes: la puntuación en el test de Aptitud Escolar (APTES) y la puntuación en el test de Aptitud Espacial (APESP).

Para la elección de estas dos variables se tienen en cuenta los resultados del análisis de componentes principales (ver Apéndice E). La variable APTES está muy correlacionada con el eje 1 ($r = 0.83$), mientras que la variable APESP está correlacionada con el eje 2 ($r = 0.64$). Una opción alternativa a esta variable es el razonamiento mecánico (DATMR).

En estos tests de aptitudes los estudiantes de nuestra muestra obtienen como media puntuaciones de un 4 ó 5 por 100 más elevadas que los estudiantes en la muestra original.

B) Otras características de los estudiantes y los centros

Incluimos variables ficticias que indican el sexo de los estudiantes (HOMBRE) y el tamaño de la población en la que residen: MENOS 10.000 = 1 si residen en una población con menos de 10.000 habitantes, MADR-BARC = 1 si residen en Madrid o Barcelona; otros tamaños de población se incluyen en el término constante de las ecuaciones de rendimiento. Solamente el 43 por 100 de nuestras observaciones corresponden a estudiantes que son mujeres. El 62 por 100 cursa estudios en centros educativos del sector público (PUBLICO = 1).

C) Características socioeconómicas

Las características socioeconómicas de los estudiantes se miden mediante un índice socioeconómico (INDICE) que representa la variabilidad conjunta (o covariabilidad) de las siguientes variables: número de hermanos, situación laboral del padre, situación laboral de la madre, profesión del padre, estabilidad-independencia del padre, estudios del padre y de la madre. Se pretende de esta forma recoger en una sola variable el efecto de una serie de factores socioeconómicos que pueden incidir en el rendimiento de los estudiantes y que por separado podrían dispersarse (ver Apéndice E).

D) Características de los compañeros de clase

La información contenida en esta muestra nos permite construir una medida de la habilidad de los compañeros de curso en cada centro educativo. A partir del análisis de componentes principales realizado (ver Apéndice E) se hallan las coordenadas de los distintos centros sobre el eje factorial 1, que se interpreta como representativo de la «habilidad» de los estudiantes.

Nos preguntamos si existe una relación positiva entre *NIVELI 1* y el aprendizaje de los estudiantes en diversas materias (en 1.º de EE.MM.), como encontraron Henderson-Miezkowski-Sauvageau (1978) en enseñanzas básicas y San Segundo (1985) en Bachiller para los estudiantes de EE.UU.

También podemos construir una medida del nivel socioeconómico de los compañeros de curso en cada centro educativo, de forma similar a como se ha hecho anteriormente para el nivel de habilidad de los estudiantes de cada centro. El análisis de correspondencias múltiples que se lleva a cabo (ver Apéndice E) proporciona las coordenadas de los distintos centros sobre el eje factorial 1, que discrimina a los estudiantes según su condición socioeconómica relativa a una serie de variables que se detallan en (C).

La muestra que utilizamos no contiene información acerca de los recursos de los centros (gasto por estudiante, ratio alumnos-profesor, instalaciones, etc.) y, lo que puede ser más importante, tampoco acerca de las características del profesor asociado a cada grupo de alumnos ni acerca de sus métodos docentes. Hanushek (1975), Murnane (1975) y Brown-Saks (1984) encontraron que algunas de estas características eran muy influyentes; por tanto, se podría mejorar la asignación de recursos educativos buscando las combinaciones óptimas (con respecto a unos ciertos objetivos) de características de profesores y alumnos.

Variable dependiente

Como variables dependientes en nuestras ecuaciones de rendimiento utilizamos los resultados obtenidos por los estudiantes en los tests de Matemáticas.

Aplicaciones (MAT. APL.), Matemáticas-Cálculo (MAT. CAL.), Comprensión Lectora (COM. LEC.) y Ortografía (ORTOG.). No utilizamos las calificaciones obtenidas por los alumnos en diversas materias por varias razones: alta tasa de no respuesta, subjetividad, heterogeneidad, etc.

4. ESTIMACION DE ECUACIONES DE RENDIMIENTO EN NIVELES Y EN DIFERENCIAS

En la Tabla 2 podemos observar las diferencias en rendimiento escolar medio entre los distintos grupos (BUP, FP y EXPE). Queremos averiguar si estas diferencias se deben a características de los estudiantes o a los programas de estudios que siguen.

En la Tabla 3 presentamos los resultados de la estimación por M.C.O. de ecuaciones de rendimiento para los 4 tests mencionados en la sección anterior, en niveles y en diferencias.

Respecto a las *características socioeconómicas* de los estudiantes observamos que el nivel socioeconómico representado por INDICE muestra una relación positiva con el rendimiento escolar solamente en Comprensión Lectora, y esta relación se hace significativa al acabar 1.º de EE.MM. (1). Si consideramos las ecuaciones en diferencias, el nivel socioeconómico de los individuos sólo muestra una relación significativa de signo positivo con el «aprendizaje» en Ortografía.

Los estudiantes que trabajan inician sus estudios secundarios con resultados significativamente inferiores en Matemáticas-Cálculo, pero al cabo de un año esta diferencia con los demás estudiantes se reduce.

En cuanto al historial académico de los estudiantes, observamos que la asistencia a *preescolar* muestra una asociación positiva significativa con el rendimiento escolar en Matemáticas-Cálculo cuando los estudiantes comienzan el 1.º curso de EE.MM. y mantienen esta ventaja al final de ese curso. Hay diferencias estadísticamente significativas entre los estudiantes que han *repetido* algún curso y aquellos que nunca lo han hecho en todos los tests. Estas diferencias persisten en EE.MM. e incluso se acentúan en Matemáticas-Aplicaciones y Matemáticas-Cálculo (véanse las ecuaciones en diferencias).

La variable que recoge los resultados de los estudiantes en el test de Aptitudes Escolares es la que muestra una relación positiva más significativa, estadísticamente, con los resultados en todos los tests. Los coeficientes más elevados corresponden a Matemáticas-Cálculo, pero en este test también se produce el mayor descen-

(1) La variable INDICE toma valores más negativos para las familias de mayor nivel socioeconómico (véase figura 1).

TABLA 2

Estadísticas descriptivas per programa de estudios

VARIABLES	BUP (N = 3.247)		FP (N = 2.354)		EXPER (N = 1.855)	
	MEDIA	DESV. S.	MEDIA	DESV. S.	MEDIA	DESV. S.
MAT. APL.	5.569	2.82	3.701	2.37	4.751	2.85
MAT. APL ₂	5.794	3.29	4.023	2.90	5.319	3.31
MAT. CALC.	17.911	5.53	13.044	5.65	14.386	5.64
MAT. CALC ₂	20.624	4.55	15.881	5.57	16.347	5.77
COMLEC	14.581	3.70	12.065	3.93	13.433	3.88
COMLEC ₂	16.339	3.45	14.190	3.79	15.527	3.71
ORTOG	37.216	5.85	32.827	7.13	34.506	7.05
ORTOG ₂	38.113	6.18	33.857	7.97	35.826	7.76
APTES	44.194	11.86	37.986	10.86	41.104	11.68
APTES ₂	53.997	12.28	45.924	11.62	50.136	12.39
APESP	5.952	4.70	5.834	5.02	5.980	4.78
APESP ₂	7.761	5.10	7.934	5.45	8.266	5.15
DATMR	38.924	7.98	38.515	8.06	38.723	8.10
DATMR ₂	43.629	7.76	43.268	7.97	44.104	7.80
PREESC.	0.893	0.31	0.828	0.37	0.884	0.32
REPET. CURS.	0.206	0.40	0.481	0.50	0.341	0.47
TRABAJA	0.361	0.48	0.407	0.49	0.390	0.48
MENOS 10.000	0.182	0.38	0.180	0.38	0.209	0.40
MADR-BARC	0.111	0.31	0.086	0.28	0.105	0.30
HOMBRE	0.481	0.50	0.665	0.47	0.575	0.49
PUBLICO	0.686	0.46	0.493	0.50	0.656	0.47
INDICE	-0.123	0.62	0.144	0.46	0.032	0.55
NIVELI 1	-0.220	0.20	0.272	0.27	0.033	0.34
SOCIOECC	44.175	4.90	38.020	4.42	41.270	5.57

TABLA 3

Ecuaciones de rendimiento escolar. N = 7.456 estudiantes

VARIABLES	NIVELES								DIFERENCIAS			
	MAT. APL.		MAT. CALC.		COMLEC.		ORTOG.		MAT. APL.	MAT. CALC.	COMLEC.	ORTOG.
	1	2	1	2	1	2	1	2				
INDICE	0.0001 (0.00)	0.0671 (1.21)	0.0359 (0.37)	-0.0258 (-0.28)	-0.1027 (-1.43)	-0.1970 (-2.88)	0.3276 (2.53)	0.0244 (0.16)	0.0669 (1.18)	-0.0618 (-0.64)	-0.0943 (-1.28)	-0.3031 (-1.91)
TRABAJA	0.0368 (0.73)	-0.0432 (-0.69)	-0.2857 (-2.64)	-0.1274 (-1.22)	0.0302 (0.37)	-0.1420 (-1.85)	-0.0198 (-0.13)	-0.1961 (-1.20)	-0.0800 (-1.26)	0.1583 (1.48)	-0.1722 (-2.10)	-0.1762 (-0.99)
PREESC	-0.0321 (-0.43)	0.1181 (1.30)	0.3750 (2.38)	0.3917 (2.58)	-0.0305 (-0.26)	0.1497 (1.34)	0.3489 (1.65)	0.0486 (0.20)	0.1502 (1.62)	0.0167 (0.10)	0.1802 (1.49)	-0.3002 (-1.16)
REP. CURS.	-0.3647 (-6.75)	-0.7269 (-10.90)	-0.4747 (-4.10)	-0.7471 (-6.71)	-0.4238 (-4.93)	-0.5294 (-6.44)	-1.4161 (-9.13)	-1.4099 (-8.04)	-0.3622 (-5.33)	-0.2723 (-2.37)	-0.1055 (-1.19)	0.0061 (0.03)
APTES	0.1110 (45.40)	0.1255 (41.60)	0.2249 (42.90)	0.1773 (35.20)	0.1340 (34.40)	0.1252 (33.60)	0.1716 (24.40)	0.1529 (19.20)	0.0145 (4.74)	-0.0476 (-9.16)	-0.0087 (-2.19)	0.0187 (-2.17)
APESP	0.0473 (8.44)	0.0413 (5.98)	0.0258 (2.15)	0.0555 (4.81)	0.0416 (4.67)	0.0311 (3.65)	-0.0231 (-1.43)	-0.0075 (-0.40)	-0.0059 (-0.85)	0.0297 (2.49)	-0.0105 (-1.14)	0.0155 (0.79)
BUP	-0.0571 (-0.77)	-0.1351 (-1.48)	0.7266 (4.59)	0.9056 (5.95)	0.3788 (3.22)	0.1608 (1.43)	1.1511 (5.43)	1.5764 (6.57)	-0.0778 (-0.83)	0.1789 (1.14)	-0.2179 (-1.80)	0.4253 (1.63)
EXPE	0.0942 (1.35)	0.3302 (3.84)	-0.6451 (-4.32)	-1.3994 (-9.76)	0.3163 (2.85)	0.3191 (3.01)	0.1776 (0.88)	0.6755 (2.99)	0.2359 (2.69)	-0.7542 (-5.10)	0.0028 (0.03)	0.4979 (2.03)

TABLA 3 (Continuación)

Ecuaciones de rendimiento escolar. N = 7.456 estudiantes

VARIABLES	NIVELES								DIFERENCIAS			
	MAT. APL.		MAT. CALC.		COMLEC.		ORTOG.		MAT. APL.	MAT. CALC.	COMLEC.	ORTOG.
	1	2	1	2	1	2	1	2				
PUBLICO	0.3333 (6.22)	0.2939 (4.44)	-0.2503 (-2.18)	-0.0990 (-0.89)	0.2273 (2.66)	0.3683 (4.52)	-0.1574 (-1.02)	0.0513 (0.29)	-0.0393 (-0.58)	0.1513 (1.33)	0.1410 (1.61)	0.2087 (1.10)
NIVEL I	-3.8274 (-24.60)	-3.0622 (-15.90)	-8.4416 (-25.30)	-7.3481 (-22.90)	-3.5141 (-14.20)	-2.0484 (-8.66)	-6.7581 (-15.10)	-4.5648 (-9.04)	0.7651 (3.91)	1.0935 (3.31)	1.4656 (5.76)	2.1933 (4.01)
SOCIOECC	-0.1247 (-12.80)	-0.0891 (-7.44)	-0.2848 (-13.70)	-0.2069 (-10.30)	-0.1393 (-8.68)	-0.0462 (-3.13)	-0.3166 (-11.30)	-0.2324 (-7.38)	0.0355 (2.92)	0.0779 (3.78)	0.0877 (5.52)	0.0842 (2.46)
MENOS 10.000	0.2064 (3.26)	0.3069 (3.93)	0.4561 (3.36)	0.4176 (3.20)	0.1919 (1.91)	0.2793 (2.90)	-0.3072 (-1.69)	0.4058 (1.97)	0.1004 (1.26)	-0.0384 (-0.30)	0.0874 (0.84)	0.7131 (3.20)
MADR-BARC	-0.1543 (-1.85)	-0.2266 (-2.20)	-0.0859 (-0.48)	-0.4225 (-2.46)	0.1815 (1.37)	0.3550 (2.80)	-0.6146 (-2.56)	-0.7325 (-2.71)	-0.0723 (-0.69)	-0.3365 (-1.90)	0.1735 (1.27)	-0.1178 (-0.40)
HOMBRE	0.2486 (4.78)	0.4604 (7.17)	-1.2206 (-10.90)	-0.0969 (-8.46)	-1.1009 (-13.30)	-1.0052 (-12.70)	-2.8227 (-18.90)	-2.8435 (-16.80)	0.2117 (3.24)	0.3137 (2.84)	0.0956 (1.12)	-0.0207 (-0.11)
CONSTANTE	4.8245	3.0067	18.3680	19.3930	13.5811	12.189	42.736	40.7440	-1.8177	1.0256	-1.3921	-1.9918
R ²	0.450	0.380	0.440	0.420	0.290	0.280	0.230	0.160	0.010	0.020	0.010	0.006
F	437.1	327.0	429.1	395.5	226.7	208.9	164.9	107.1	8.2	11.6	5.4	3.4

- Estimaciones por mínimos cuadrados ordinarios.
- t, estadísticos en paréntesis.

so en el coeficiente de APTES tras el 1.º curso de EE.MM. El test de Aptitud Espacial igualmente muestra una relación positiva y estadísticamente significativa con los resultados en todos los tests excepto en Ortografía (no significativa) (2).

Observamos que los alumnos que cursan *BUP* comienzan sus estudios con ventajas estadísticamente significativas en todos los tests, excepto en Matemáticas-Aplicadas. Después de un año, las diferencias a su favor se mantienen, y aumentan ligeramente en el caso de la Ortografía. Por el contrario, los alumnos que cursan *Enseñanzas Experimentales* comienzan sus estudios con ventaja en Comprensión Lectora (en relación con los estudiantes que inician FP) y con clara desventaja en Matemáticas-Cálculo (3). Después del primer curso de EE.MM. mantienen la ventaja en Comprensión Lectora, avanzan considerablemente en Ortografía y en Matemáticas-Aplicaciones y empeoran su situación en Matemáticas-Cálculo (en este test sus resultados son siempre significativamente inferiores a los obtenidos por los alumnos de BUP y por los de FP).

Los alumnos que se matriculan en centros *públicos* parten con ventajas estadísticas significativas en Matemáticas Aplicadas y en Comprensión Lectora y mantienen estas ventajas tras su primer año en EE.MM. En Matemáticas-Cálculo eliminan su desventaja inicial de un año.

VARIABLES CONTEXTUALES

Los estudiantes que se matriculan en centros con estudiantes de habilidad elevada o de alto nivel socioeconómico comienzan EE.MM. con significativas diferencias positivas en todos los tests, pero pierden parte de estas ventajas iniciales (4).

Los alumnos que se matriculan en centros escolares situados en poblaciones de menos de 10.000 habitantes presentan ventajas en sus niveles de rendimiento en Matemáticas (Aplicaciones y Cálculo) y en Comprensión Lectora; en Ortografía presentan esta ventaja al finalizar el 1.º curso de EE.MM., al mismo tiempo que mantienen las diferencias a su favor en los otros tests.

Los estudiantes de Madrid y Barcelona muestran niveles de rendimiento significativamente inferiores en todos los tests excepto en Comprensión Lectora.

(2) En el Apéndice A comprobamos que estas medidas de habilidad no crean serios problemas de errores en variables.

(3) En Modrego y San Segundo (1987) estimamos ecuaciones separadas para los tres grupos (BUP, FP y EXPE) y las conclusiones no varían.

(4) En Modrego y San Segundo (1987) comprobamos que estos efectos son cóncavos para la habilidad (NIVELI 1) y convexos para SOCIOECC, como se ha observado en estudios con datos norteamericanos.

Los hombres tienen mejores resultados que las mujeres solamente en Matemáticas Aplicadas; en 1.º de EE.MM. mejoran significativamente su rendimiento en Matemáticas Aplicadas y en Matemáticas-Cálculo.

CONCLUSIONES

En cuanto a la significación global de cada ecuación, observamos que los mejores ajustes corresponden a las ecuaciones de matemáticas en niveles ($R^2 = 0.4$), y los peores a todas las ecuaciones en diferencias. Nuestra especificación claramente no explica la evolución del rendimiento escolar en 1.º curso de EE.MM. Nuestras *ecuaciones en diferencias* no nos dicen por qué aprenden los individuos en ninguna de estas cuatro materias (5). Dada la escasa significación de estas ecuaciones y la difícil interpretación de las estimaciones obtenidas, debemos concentrarnos en los resultados correspondientes a las *ecuaciones en niveles*, aunque comentaremos los resultados que producen las ecuaciones en diferencias.

De acuerdo con los resultados contenidos en la Tabla 3, el comportamiento del rendimiento escolar en Matemáticas (especialmente en Matemáticas-Cálculo) es diferente al comportamiento del rendimiento en Comprensión Lectora y en Ortografía.

Así, el hecho de haber asistido a educación preescolar sólo mantiene una relación, al iniciar y al acabar 1.º curso de EE.MM., con el rendimiento en Matemáticas-Cálculo. El hecho de haber repetido algún curso está asociado negativamente con los niveles alcanzados en todos los tests; en las ecuaciones en diferencias observamos que sólo mantiene una relación negativa con el aprendizaje en Matemáticas en EE.MM.

APTES tiene el coeficiente positivo más elevado en Matemáticas-Cálculo y está asociado positivamente con el aprendizaje en Matemáticas-Aplicaciones. APESP tiene una relación positiva con el aprendizaje en Matemáticas-Cálculo. Por el contrario, el índice socioeconómico sólo presenta una relación positiva significativa con el rendimiento al final del 1.º curso en Comprensión Lectora y con el aprendizaje en Ortografía. Parece que los factores de habilidad están más asociados al rendimiento en Matemáticas y los factores socioeconómicos muestran alguna relación con el rendimiento verbal.

Probablemente el resultado más llamativo de esta tabla sea la relación negativa entre los resultados (y el aprendizaje) en Matemáticas-Cálculo y la participación en el Programa Experimental de EE.MM.

Otro resultado que encontramos llamativo son los coeficientes «favorables a los centros públicos» que contiene la Tabla 3. Los estudiantes que se matriculan

(5) Alternativamente, las variaciones en los resultados en los tests desde 1984 a 1985 no miden aprendizaje.

en centros públicos presentan resultados mejores en Matemáticas-Aplicaciones y en Comprensión Lectora que los estudiantes de los centros privados; tras hacer el 1.º curso de las EE.MM. mantienen estas ventajas iniciales y pierden su desventaja de partida en Matemáticas-Cálculo. Estos resultados contradicen las afirmaciones habituales acerca de los mejores resultados académicos de los centros privados de enseñanza.

De acuerdo con los datos de la Tabla 3, el balance de la Reforma para los estudiantes de nuestra muestra es positivo en Ortografía y en Matemáticas-Aplicaciones, pero negativo en Matemáticas-Cálculo.

En la próxima sección evaluamos la Reforma comparando los estudiantes experimentales con estudiantes de BUP y FP que asisten a sus mismos centros.

5. EVALUACION DE LA REFORMA CON DATOS EXPERIMENTALES

En las ciencias sociales se concede gran importancia a la posibilidad de evaluar programas (sanitarios, educativos, de empleo, etc.) utilizando datos provenientes de un experimento.

En un problema de evaluación de programas de Formación Profesional, Robert Lalonde (1986) ha comparado los efectos estimados con datos experimentales y las estimaciones obtenidas comparando individuos participantes en un programa con individuos «similares» de la población (el equivalente a nuestro grupo de control externo). Sus resultados indican que las estimaciones difieren considerablemente según el grupo de comparación elegido y la técnica estadística empleada para estimar los efectos del programa. En general, resulta difícil replicar los resultados correspondientes a los datos experimentales cuando se construyen grupos de control ajenos al experimento. Vamos a estudiar la importancia de este problema para la evaluación de la Reforma de las EE.MM.

En esta sección evaluamos el rendimiento obtenido por los estudiantes experimentales ($n = 1.855$) comparándolo con el rendimiento obtenido por el grupo de control interno, es decir, por los 1.293 estudiantes de nuestra muestra que cursaban BUP o FP en los mismos centros experimentales. En la Tabla 4 aparecen las ecuaciones de rendimiento escolar estimadas con estos datos experimentales ($n = 3.148$). Si comparamos los coeficientes recogidos en esta tabla con los de la Tabla 3 ($n = 7.456$), vemos que apenas hay variaciones en el signo o en la magnitud de la mayoría de los coeficientes; sólo en algunos casos se reduce la significación de los coeficientes al utilizar una muestra de menor tamaño.

Los cambios más destacados se producen en las variables «dummy» que recogen el tamaño de la población (MENOS 10.000 y MADR-BARC). Así, residir en Madrid o en Barcelona está ahora significativamente asociado con bajo rendimiento en Ortografía y con una caída en el rendimiento en Matemáticas-Cálculo.

TABLA 4

Ecuaciones de rendimiento escolar para los alumnos de centros experimentales. N = 3.148

	NIVELES							
	MAT. APL.		MAT. CALC.		COMLEC.		ORTOG.	
	1	2	1	2	1	2	1	2
INDICE	0.08 (1.22)	0.06 (0.54)	0.02 (0.00)	0.06 (0.31)	-0.17 (-1.50)	-0.21 (-1.90)	0.03 (0.17)	-0.05 (-0.30)
TRABAJA	0.03 (0.44)	-0.12 (-1.20)	-0.12 (-0.70)	-0.49 (-2.80)	-0.02 (0.00)	-0.06 (-0.50)	0.08 (0.31)	-0.21 (0.70)
PREESC	-0.09 (-0.77)	0.00 (0.00)	0.29 (1.14)	0.56 (5.20)	-0.00 (0.00)	0.39 (2.23)	0.48 (1.45)	-0.01 (0.00)
REPT. CURS.	-0.32 (-3.8)	-0.72 (-6.8)	-0.46 (-2.5)	-0.63 (-3.4)	-0.48 (-3.6)	-0.52 (-4.1)	-1.67 (-6.9)	-1.84 (-6.5)
APTES	0.11 (29.0)	0.12 (25.1)	0.22 (26.2)	0.17 (21.1)	0.13 (21.2)	0.12 (21.1)	0.18 (16.2)	0.15 (11.4)
APESP	0.060 (6.40)	0.050 (4.34)	0.026 (1.34)	0.081 (4.17)	0.030 (1.90)	0.030 (2.20)	-0.040 (-1.45)	0.005 (0.00)
BUP	0.04 (0.31)	0.12 (0.70)	1.59 (5.44)	1.88 (1.37)	0.49 (2.30)	0.23 (1.14)	1.92 (4.95)	2.58 (5.70)
EXPE	0.14 (1.22)	0.26 (1.87)	0.12 (0.54)	-0.83 (-3.50)	0.26 (1.48)	0.32 (1.92)	0.65 (2.05)	1.33 (3.60)
PUBLICO	0.41 (4.50)	0.37 (5.11)	-0.64 (-3.30)	-0.71 (-3.60)	0.28 (1.92)	0.40 (2.94)	0.32 (1.22)	-0.14 (-0.40)
NIVELI 1	-3.48 (-14.6)	-3.13 (-10.6)	-6.11 (-11.9)	-6.77 (-13.3)	3.06 (-8.1)	-2.20 (-6.1)	-6.02 (-8.8)	-5.01 (-6.3)
SOCIOECC	-0.09 (-6.8)	-0.08 (-4.6)	-0.16 (-5.3)	-0.13 (-4.5)	-0.10 (-4.3)	-0.04 (-2.1)	-0.24 (-5.9)	-0.22 (-4.8)
MENOS 10.000	0.18 (1.73)	0.42 (3.33)	1.02 (4.60)	0.29 (1.34)	0.58 (3.62)	0.50 (3.22)	-0.32 (-1.20)	0.40 (1.18)
MADR-BARC	-0.16 (-1.30)	0.06 (0.31)	0.12 (0.43)	-0.78 (-2.90)	0.32 (1.61)	0.72 (3.80)	-1.33 (-3.60)	-1.3 (-3.10)
HOMBRE	0.30 (3.63)	0.52 (5.10)	-1.22 (-6.90)	-0.92 (-5.20)	-1.06 (-8.20)	-1.10 (-8.90)	-2.7 (-11.60)	-2.68 (-9.70)

Se mantienen las relaciones positivas significativas entre nuestras medidas de habilidad (APTES y APESP) y el rendimiento escolar. Los estudiantes de BUP ahora presentan mejores resultados que los estudiantes de FP en todos los tests (diferencias significativas sólo se mantienen en Ortografía) y que los estudiantes de Enseñanzas Experimentales en todos los tests excepto en Matemáticas-Aplicaciones. En la muestra total (Tabla 3) BUP no aventajaba a FP en Matemáticas-Aplicaciones ni a Experimentales, Matemáticas-Aplicaciones y en Comprensión Lectora.

Los alumnos de 1.º curso de *Enseñanzas Experimentales* empeoran su situación en Matemáticas-Cálculo, como muestra la Tabla 3, y mejoran su situación en Ortografía. Pero su avance en Matemáticas-Aplicaciones parece comparable al de los demás estudiantes de los mismos centros, y tampoco manifiestan ahora una ventaja muy significativa en Comprensión Lectora (antes aventajaban a FP y a BUP en el conjunto de la muestra).

TABLA 4 (Continuación)

Ecuaciones de rendimiento escolar para los alumnos de centros experimentales. N = 3.148

	NIVELES							
	MAT. APL.		MAT. CALC.		COMLEC.		ORTOG.	
	1	2	1	2	1	2	1	2
CONSTANTE	3.48	2.90	12.78	16.30	12.20	12.00	38.40	40.30
R ²	0.42	0.35	0.37	0.38	0.26	0.26	0.23	0.15
F	160.1	120.7	132.9	136.5	78.4	80.0	66.9	41.0

- Estimaciones por mínimos cuadrados ordinarios.
- t, estadísticos en paréntesis.

6. CONCLUSIONES

En nuestro trabajo hemos utilizado información referente a tres grupos de alumnos: los que cursan EE.MM. Experimentales (EXPE; n = 1.855), los que cursan BUP y FP en centros experimentales (Control interno, CONT. I.; n = 1.293) y los que cursan BUP y FP en centros no experimentales (Control externo, CONT. E.; n = 4.308). Nuestras estimaciones de los efectos de la Reforma de las EE.MM. son las mismas cuando comparamos EXPE con CONT. I. (Tabla 4) que cuando lo hacemos con CONT. E. (Tabla 3), y esta similitud refuerza nuestra confianza en los resultados obtenidos.

Nuestras conclusiones básicas son las siguientes:

- A. Las medidas de habilidad que utilizamos muestran una relación significativa con el rendimiento en todas las materias, especialmente en Matemáticas. Las medidas de status socioeconómico familiar sólo presentan alguna asociación con el rendimiento verbal.
- B. En cuanto al historial académico de los estudiantes, se observa una asociación positiva entre la asistencia a preescolar y el rendimiento en Matemáticas-Cálculo, y persisten las diferencias en EE.MM. entre los estudiantes que han repetido algún curso y aquellos que nunca lo han hecho.
- C. Los alumnos que se matriculan en centros públicos inician EE.MM. con ventajas significativas en Matemáticas-Aplicadas y en Comprensión Lectora y mantienen estas ventajas tras el 1.º año en EE.MM. Nuestros resultados contradicen las afirmaciones habituales acerca de los mejores resultados académicos en los centros privados de enseñanza, en comparación con los públicos.
- D. Los alumnos de enseñanzas experimentales mejoran significativamente su rendimiento en Matemáticas-Aplicaciones y Ortografía y empeoran su situación en Matemáticas-Cálculo en relación con los estudiantes de BUP y FP.

APENDICE A: MEDIDAS DE HABILIDAD DE LOS ESTUDIANTES.
¿ERRORES EN VARIABLES?

En las ecuaciones de rendimiento que estimamos en el texto, las medidas de las aptitudes de los estudiantes jugaban un papel muy importante: mostraban una relación significativa con el rendimiento escolar en todos los tests. Hemos argumentado que estas variables «medían» la habilidad de los estudiantes, pero sabemos que sólo son indicadores imperfectos de habilidad. El introducirlos puede plantear varios problemas en nuestras ecuaciones y deberemos emplear diferentes métodos de estimación para resolverlos. Veamos algunos de estos problemas.

A. Es probable que APTES y APESP midan habilidades de los individuos con errores. Supongamos, por ejemplo, que APTES mide la habilidad «de tipo 1» más un error aleatorio y que APESP tiene una relación similar con la habilidad de tipo 2, es decir,

$$\text{APTES}_j = \text{HAB } 1_j + \xi_j$$

$$\text{APESP}_j = \text{HAB } 2_j + \eta_j$$

En este caso, nuestros estimadores de mínimos cuadrados ordinarios (MCO) de los parámetros de las ecuaciones de rendimiento son inconsistentes. Para resolver este problema podemos pensar en estimar las ecuaciones de rendimiento por el método de Variables Instrumentales (V.I.). Necesitamos encontrar las variables Z_{1j} , Z_{2j} muy relacionadas con HAB_1 y HAB_2 , respectivamente, e independientes de ξ_j y de η_j .

Algunas soluciones propuestas por Griliches (1977) en problemas análogos son las siguientes:

- Se pueden utilizar las características socioeconómicas de los estudiantes como variables instrumentales; aunque esta solución no es muy adecuada para nuestras ecuaciones de rendimiento, que ya incluyen estas características como variables exógenas.
- Si se dispone de medidas adicionales de estas habilidades, se pueden utilizar como variables instrumentales de APTES y APESP; por ej., si tenemos

$$\text{APTES}_{jt-1} = \text{HAB } 1_j + \xi_{jt-1}$$

$$\text{APESP}_{jt-1} = \text{HAB } 2_j + \eta_{jt-1}$$

$$\text{APTES}_{jt} = \text{HAB } 1_j + \xi_{jt}$$

$$\text{APESP}_{jt} = \text{HAB } 2_j + \eta_{jt}$$

y si los errores de medida son independientes, podemos usar $APTES_t$ y $APESP_t$ como instrumentos para $APTES_{t-1}$ y $APESP_{t-1}$.

Esta solución parece, en principio, la más adecuada para nuestro problema, pero plantea dificultades de computación con los paquetes estadísticos disponibles en la UPV. Si los errores de medida son importantes, puede que sea preferible excluir $APTES$ y $APESP$ de las ecuaciones de rendimiento para evitar el problema de inconsistencia que causa su introducción.

B. Es posible que las medidas de rendimiento y las de aptitud estén afectadas por componentes aleatorios comunes; es decir, puede que el *rendimiento* y la *aptitud* sean medidas de resultados obtenidos por el individuo afectadas por su capacidad intelectual, motivación, etc. En este caso nuestras estimaciones de MCO también serán inconsistentes y, de nuevo, necesitaremos encontrar variables instrumentales para $APTES$ y $APESP$, es decir, variables muy correlacionadas con estas aptitudes e independientes de la motivación, la inteligencia, etc.

Nuevamente nos encontramos con el problema de identificar variables exógenas no incluidas ya en nuestras ecuaciones de rendimiento. La alternativa que vamos a considerar consiste en excluir $APTES$ y $APESP$ de las ecuaciones de rendimiento de la sección anterior para observar los efectos de su exclusión sobre los demás coeficientes.

En la Tabla 5 presentamos estas nuevas estimaciones en niveles y en diferencias para los cuatro tests disponibles.

Observamos que el INDICE socioeconómico muestra una relación positiva constante con el rendimiento en Matemáticas-Aplicaciones, Matemáticas-Cálculo y en Comprensión Lectora y que está asociado al aprendizaje en Ortografía.

Seguimos obteniendo coeficientes positivos para la variable que indica que el estudiante asistió a Preescolar y coeficientes negativos para la variable que indica si el estudiante ha repetido curso alguna vez.

De nuevo observamos que los individuos que se matriculan en BUP tienen un rendimiento elevado en todos los tests, excepto en Matemáticas-Aplicaciones, y que los que comienzan Enseñanzas Experimentales tienen ventajas (sobre los de FP) en Comprensión Lectora, y un rendimiento muy bajo en Matemáticas-Cálculo.

Comparando los coeficientes de la Tabla 3 con los que hemos obtenido al excluir $APTES$ y $APESP$ de las ecuaciones podemos concluir que:

1. El mayor cambio se produce en los coeficientes del INDICE, que mide el nivel socioeconómico del estudiante. Al no incluir medidas de habilidad, se hacen más significativos y aumentan en valor absoluto.

Parece que existe una asociación positiva entre niveles socioeconómicos y habilidades (ya que en la Tabla 5 no se incluyen medidas de habilidad, su relación con el rendimiento pasa a ser recogida por INDICE).

2. Los otros cambios perceptibles corresponden a los coeficientes de NIVELI 1 y de SOCIOECC.

El índice de habilidad del centro (NIVELI 1) también parece tener una mayor relación con el rendimiento cuando no incluimos medidas de habilidad del individuo en las ecuaciones. De nuevo esperamos que exista una relación positiva entre la habilidad del estudiante y la de sus compañeros que explique este resultado.

El índice que recoge el nivel socioeconómico de los alumnos de 1.º curso en el centro (SOCIOECC) tiene coeficientes menores cuando no incluimos APTES y APESP en las ecuaciones. Este resultado no parece de fácil interpretación.

En vista de que la introducción de nuestras medidas de habilidad en las ecuaciones de rendimiento no parece sesgar los coeficientes estimados y de que mejora el ajuste global de las ecuaciones, concluimos que los problemas de errores en las variables no parecen ser serios en este caso, y elegimos mantener APTES y APESP en las ecuaciones de rendimiento escolar.

TABLA 5

Ecuaciones de rendimiento escolar sin medidas de habilidad. N = 7.456

VARIABLES	NIVELES								DIFERENCIAS			
	MAT. APL.		MAT. CALC.		COMLEC.		ORTOG.		MAT. APL.	MAT. CALC.	COMLEC.	ORTOG.
	1	2	1	2	1	2	1	2				
INDICE	-0.16 (-2.98)	-0.11 (-1.67)	-0.26 (-2.40)	-0.27 (-2.60)	-0.29 (-3.70)	-0.37 (-4.90)	0.11 (0.83)	-0.17 (-1.10)	0.0500 (0.88)	-0.0081 (-0.10)	-0.0800 (-1.10)	-0.2830 (-1.70)
TRABAJA	0.05 (0.90)	-0.03 (-0.38)	-0.257 (-2.10)	-0.11 (-0.92)	0.05 (0.54)	-0.12 (-1.50)	0.02 (0.00)	-0.18 (1.10)	-0.0780 (-1.25)	0.1520 (1.41)	-0.1730 (-2.10)	-0.1788 (-1.00)
PREESC	0.12 (1.45)	0.29 (2.84)	0.684 (3.83)	0.64 (3.84)	0.15 (1.22)	0.32 (2.66)	0.58 (2.64)	0.26 (1.10)	0.1697 (1.83)	-0.0460 (-0.30)	0.1680 (1.40)	-0.3247 (-1.20)
REPET. CURS.	-0.66 (-10.6)	-1.06 (-14.1)	-1.04 (-8.0)	-1.21 (-9.9)	-0.77 (-8.3)	-0.85 (-9.6)	-1.83 (-11.3)	-1.78 (-9.9)	-0.3953 (-5.85)	-0.1691 (-1.40)	-0.0789 (-0.90)	0.0450 (0.24)
BUP	-0.13 (-1.60)	-0.21 (-2.10)	0.63 (3.49)	0.80 (4.80)	0.30 (2.32)	0.09 (0.77)	1.11 (5.00)	1.53 (6.21)	-0.0784 (-0.84)	0.1729 (1.10)	-0.2065 (-1.70)	0.4200 (1.61)
EXPE	0.05 (0.63)	0.28 (2.89)	-0.72 (-4.20)	-1.47 (-9.33)	0.27 (2.20)	0.27 (2.40)	0.13 (0.63)	0.63 (2.72)	0.2330 (2.66)	-0.7467 (-4.80)	-	0.5000 (2.04)
PUBLICO	0.32 (5.20)	0.29 (3.83)	-0.25 (-1.90)	-0.11 (-0.90)	0.22 (2.36)	0.36 (4.11)	-0.14 (-0.90)	0.06 (0.31)	-0.0368 (-0.50)	0.1399 (1.22)	0.1439 (1.64)	0.2030 (1.07)
NIVELI 1	-4.27 (-23.8)	-3.53 (-16.3)	-9.12 (-24.3)	-8.00 (-22.9)	-4.00 (-15.0)	-2.48 (-9.7)	-7.13 (-15.4)	-4.94 (-9.6)	0.7466 (3.84)	1.1200 (3.40)	1.5200 (6.02)	2.1900 (4.03)

TABLA 5 (Continuación)

Ecuaciones de rendimiento escolar sin medidas de habilidad. N = 7.456

VARIABLES	NIVELES								DIFERENCIAS			
	MAT. APL.		MAT. CALC.		COMLEC.		ORTOG.		MAT. APL.	MAT. CALC.	COMLEC.	ORTOG.
	1	2	1	2	1	2	1	2				
SOCIOECC	-0.04 (-3.98)	0.28 (0.20)	-0.118 (-5.10)	-0.077 (-3.60)	-0.036 (-2.20)	0.05 (2.93)	-0.19 (-6.50)	-0.12 (-3.70)	0.047 (3.91)	0.041 (2.00)	0.082 (5.25)	0.069 (2.07)
MENOS 10.000	0.20 (2.80)	0.31 (3.50)	0.48 (3.13)	0.42 (2.96)	0.20 (1.78)	0.29 (2.73)	-0.27 (-1.40)	0.43 (2.04)	0.1050 (1.32)	-	0.0902 (0.87)	-
MADR-BARC	-0.13 (-1.30)	-0.2 (-1.70)	-0.053 (-0.26)	-0.39 (-2.10)	0.21 (1.41)	0.38 (2.73)	-0.50 (-2.40)	-0.72 (-2.60)	-0.072 (-0.68)	-	0.1701 (1.25)	-
HOMBRE	0.66 (11.4)	0.89 (12.8)	-0.60 (-4.9)	-0.31 (-2.7)	-0.65 (-7.5)	-0.60 (-7.4)	-2.46 (-16.6)	-2.49 (-15.0)	0.2307 (3.68)	0.2807 (2.64)	0.0440 (0.53)	-0.0230 (-0.14)
CONSTANTE	6.13	4.40	20.50	21.30	15.05	13.50	44.06	42.00	-	0.8293	-1.5500	-2.041
R ²	0.26	0.21	0.29	0.31	0.16	0.15	0.17	0.12	0.012	0.010	0.011	0.005
F	218.3	160.9	255.3	280.5	121.79	112.1	128.5	86.7	7.62	6.53	5.67	3.64

- Estimaciones por mínimos cuadrados ordinarios.
 - t, estadísticos en paréntesis.

APENDICE ESTADISTICO (E)

E.1. *Introducción*

En este apéndice se describen los procedimientos empleados para obtener los índices socioeconómicos (variables INDICE y SOCIOECC) y de habilidad (variable NIVELI 1) a partir de los datos recogidos en las encuestas del CIDE.

E.2. *Índice socioeconómico*

El método utilizado para construir el índice socioeconómico es el análisis factorial de correspondencias múltiples. Con este método es posible obtener dos tipos de variables:

- una variable individual, INDICE, que está fuertemente correlacionada con variables que, se supone, están a su vez relacionadas con el nivel socioeconómico de cada estudiante;
- una variable asociada a cada centro, SOCIOECC, que recoge el nivel socioeconómico conjunto de los alumnos que cursan estudios en cada uno de los centros.

E.2.1. *Variables*

Las variables activas que intervienen en el análisis se relacionan en función de su supuesta relevancia inicial en la determinación del nivel socioeconómico de los estudiantes. Son las siguientes: número de hermanos (HERMA2), situación laboral del padre (SILAPA2), situación laboral de la madre (SILAMA2), profesión del padre (PROFPAD2), estabilidad-independencia del padre (ESTALAP2), estudios del padre (ESTPAD2) y estudios de la madre (ESTMAD2). Todas estas variables están relacionadas con el nivel socioeconómico de los estudiantes. Además, la fuerte relación que existe entre ellas se pondrá de manifiesto en el alto porcentaje de inercia o varianza recogida por el primer eje que se obtenga como resultado de esta relación. Esto, que puede representar un problema, si lo que se pretende es describir un conjunto de observaciones, es una ventaja en este caso en el que se trata de construir un índice representativo del nivel socioeconómico de cada estudiante. De esta forma, si se llega a encontrar un eje que recoja la covariabilidad de estas

variables, la posición de cada individuo con relación a dicho eje se considerará un indicador de su nivel socioeconómico. Lo mismo se puede decir en cuanto a la posición de cada centro.

Aquellas variables que tienen un efectivo menor o igual que 20 se consideran, a efectos del análisis, modalidades suplementarias. Para que el número de individuos sobre los que se efectúa el análisis se mantenga constante, el programa establece una asignación aleatoria de los individuos que tienen la modalidad suprimida como activa entre el resto de las modalidades.

E.2.2. Valores propios

En la Tabla E.1. aparecen los valores propios resultantes del análisis que representan la dispersión o inercia de la nube de puntos sobre cada uno de los ejes factoriales. La tasa de inercia o porcentaje de información correspondiente a los dos primeros ejes representa el 9.32 por 100 de la inercia o varianza total de la nube de puntos. A primera vista, estos valores relativamente pequeños pueden dar una idea muy pesimista de la información recogida por estos ejes; sin embargo, son valores habituales en el análisis de correspondencias múltiples. Para solventar este problema Benzecri (1979) propone calcular las tasas de información sobre los valores propios transformados $(\lambda_{\alpha} - 1/Q)^2$, donde Q es el número de variables, no reteniendo más que aquellos $\lambda_{\alpha} > 1/Q$. Los porcentajes de información obtenidos mediante esta transformación (Tabla E.2.) son muy superiores. Queda así puesta de manifiesto la importancia del primer eje, a partir del cual se pretende construir la variable artificial índice socioeconómico.

E.2.3. Coordenadas, contribuciones absolutas y relativas.

Representación gráfica

Por definición, las contribuciones absolutas representan la contribución de cada modalidad a la dispersión explicada por cada uno de los ejes factoriales. Las contribuciones relativas miden la parte de la dispersión de cada fila explicada por un eje factorial. Para la correcta interpretación de los resultados ha de tenerse en cuenta lo siguiente:

- la distancia de una modalidad al centro de gravedad es función decreciente de su efectivo total,
- la contribución de una modalidad a la inercia total es una función decreciente de su efectivo total,
- la inercia debida a una variable es función creciente del número de modalidades de respuesta.

Las variables profesión del padre, estudios del padre y estudios de la madre contribuyen de forma destacada a la inercia explicada por el primer eje. Ellas representan el 81.3 por 100 de la inercia de dicho eje. Las características más sobresalientes del análisis están reflejadas en la Figura 1, que representa las proyecciones de las distintas modalidades, activas y suplementarias, sobre los planos formados por los dos primeros ejes factoriales.

En la Figura 1 se aprecia la ordenación, de izquierda a derecha, de las variables profesión del padre, estudios del padre y estudios de la madre, que justifican la interpretación del eje 1 como representativo del nivel socioeconómico del estudiante. En este plano se aprecia una forma parabólica de la nube de puntos debida al *efecto Guttman* que aparece cuando la nube de puntos puede reducirse a una sola dimensión.

Las coordenadas de los centros respecto al eje factorial proporcionan una medida del nivel socioeconómico de los alumnos que cursan estudios en cada uno de ellos.

E.3. Nivel de habilidad

Para medir el nivel de habilidad de cada centro (variable NIVELI 1) se utiliza el análisis factorial de componentes principales, cuya aplicación se describe en los apartados siguientes.

E.3.1. Variables

Las variables activas a partir de las cuales se efectúa el análisis corresponden a los resultados de las pruebas de los siguientes tests en la primera aplicación: Puntuación Global Aptitudes Escolares (APTES), Razonamiento Mecánico (DATMR), Aptitud Espacial (APESP), Matemáticas-Cálculo (MTCAL), Matemáticas-Aplicadas (MTAPL), Ortografía (ORTOG) y Comprensión Lectora (COMLEC). Como variables ilustrativas se consideran los valores de los mismos tests en la segunda observación. Los principales estadísticos de estas variables figuran en la Tabla 1 del texto. En la Tabla E.3 aparece la matriz de correlaciones a partir de la cual se efectúa el análisis; en la lectura de esta tabla se aprecia la covariabilidad positiva entre todas las variables, más fuerte entre algunas de ellas, lo que quedará reflejado en los resultados posteriores.

E.3.2. Valores propios

El primer resultado numérico interpretable lo constituye la lista de los valores propios y porcentajes de varianza que proporciona la Tabla E.4. Los primeros por-

FIGURA 1
PLAN DE PROJECTION DES 62 PCINTS SUR LES ASES 2 ET 3
AXE 2/HORIZONTAL AXE 3/VERTICAL

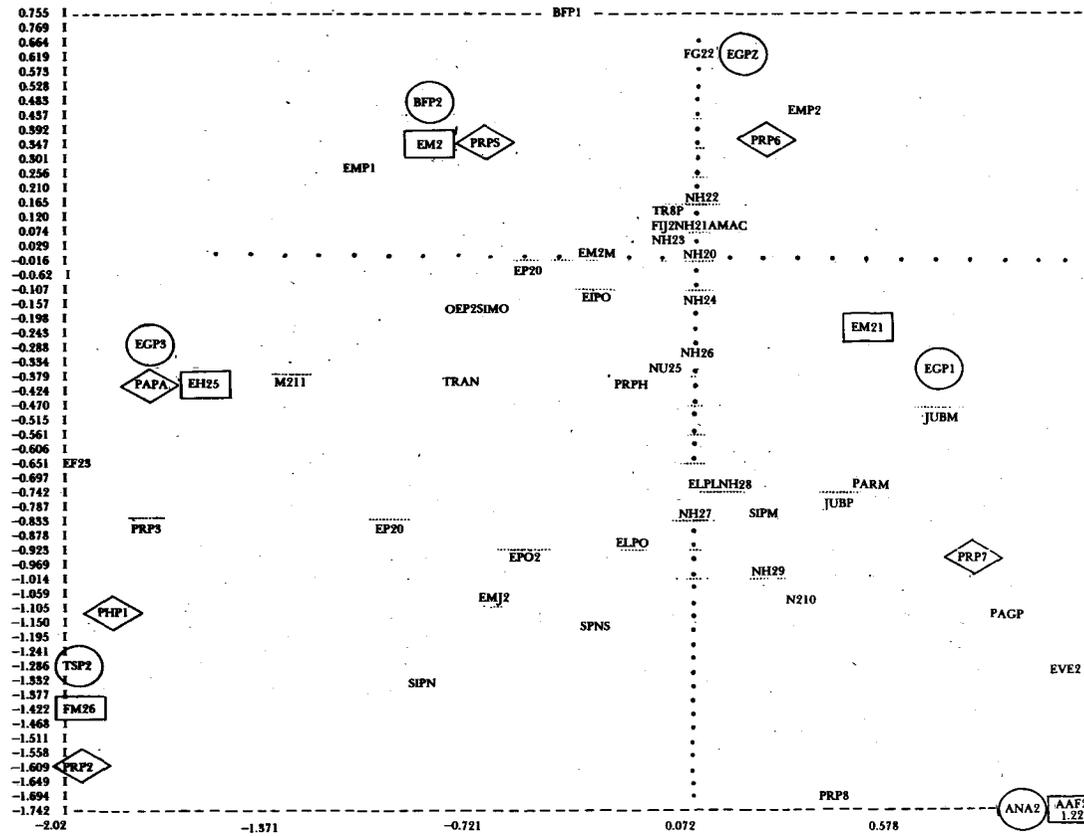


TABLA E.3.
Matriz de correlaciones

	APTES	DATMR	APESP	MTCAL	MTAPL	ORTOG	COMLEC
APTES	1.00	0.39	0.35	0.56	0.59	0.34	0.47
DATMR	0.35	1.00	0.47	0.21	0.33	0.01	0.21
APESP	0.35	0.47	1.00	0.19	0.29	0.05	0.17
MTCAL	0.56	0.21	0.19	1.00	0.63	0.39	0.39
MTAPL	0.59	0.33	0.29	0.63	1.00	0.32	0.40
ORTOG	0.34	0.01	0.05	0.39	0.32	1.00	0.33
COMLEC	0.47	0.21	0.17	0.39	0.40	0.33	1.00

TABLA E.4.
Valores propios

	V. PROPIOS	%	% ACUMUL.
1	3.12800800	44.69	44.69
2	1.27421700	18.20	62.89
3	0.68273700	9.75	72.64
4	0.65533290	9.36	82.00
5	0.51097040	7.30	89.30
6	0.39875140	5.70	95.00
7	0.34906600	5.00	100.00

centajes son bastante elevados (recogen el 62.89 por 100 de la varianza), lo cual es debido, por una parte, a la pequeña dimensión de la matriz de correlaciones, y por otra, a la concentración de la nube en un subespacio de dos dimensiones.

E.3.3. Coordenadas, contribuciones absolutas y relativas. Representación gráfica

Como se sabe, las coordenadas de los puntos variables son menores que la unidad en valor absoluto y su signo sobre cada eje depende de la orientación de éstos, que a su vez depende del algoritmo de diagonalización utilizado.

En la Tabla E.5 se observa cómo las coordenadas de todas las variables sobre el primer eje factorial son positivas; es decir, todas ellas están a un mismo lado del origen. Esto se debe a la correlación positiva existente entre todas las variables. Esta característica aparece a menudo sobre el primer eje, que pasa a denominarse *factor de talla*. En este caso, el factor de talla contrapone los estudiantes «más hábiles» a los «menos hábiles».

En la Figura 2 aparece la representación gráfica de los puntos variables sobre los dos primeros ejes. El primer eje factorial se puede interpretar como una medida de lo que se podría denominar «habilidad de los estudiantes».

Las abscisas de los centros de educación sobre el primer eje factorial aparecen como una medida o índice del nivel de habilidad de los alumnos que pertenecen a cada uno de ellos. Por construcción, este índice asegura una dispersión máxima sobre el conjunto de individuos, por lo que le da el carácter de índice discriminatorio entre centros. Las posiciones de los puntos representativos de los centros en la Figura 3 son a este respecto elocuentes de las diferencias existentes en el nivel de habilidad de los estudiantes de los distintos centros.

TABLA E.5.

Coordenadas

NOMS	E TYPE	COORDONNEES						PROJECTION ANCIENS AXES UNITE (CARRE-CONTRIBUTION ABSOLUET)						CORRELATION VARIABLE-FACTEUR (CARRE-CONTRIBUTION RELATIVE)					
		F1	F2	F3	F4	F5	F6	F1	F2	F3	F4	F5	F6	F1	F2	F3	F4	F5	F6
APTE	11.813	-0.83	0.01	0.05	-0.08	0.00	0.00	-0.47	0.01	0.06	-0.10	0.00	0.00	-0.83	0.01	0.05	-0.08	0.00	0.00
DATM	8.028	-0.53	0.66	-0.06	-0.04	0.00	0.00	-0.30	0.58	-0.07	-0.05	0.00	0.00	-0.53	0.66	-0.06	-0.04	0.00	0.00
APESP	4.826	-0.50	0.64	-0.22	0.32	0.00	0.00	-0.28	0.57	-0.27	0.39	0.00	0.00	-0.50	0.64	-0.22	0.32	0.00	0.00
MTCAL	6.010	-0.77	-0.27	0.38	0.05	0.00	0.00	-0.43	-0.24	0.47	0.06	0.00	0.00	-0.77	-0.27	0.38	0.05	0.00	0.00
MTAPL	2.816	-0.81	-0.07	0.36	-0.02	0.00	0.00	-0.40	-0.06	0.43	-0.02	0.00	0.00	-0.81	-0.07	0.36	-0.02	0.00	0.00
ORTOG	6.854	-0.51	-0.56	-0.37	0.50	0.00	0.00	-0.29	-0.49	-0.45	0.62	0.00	0.00	-0.51	-0.56	-0.37	0.50	0.00	0.00
COMLEC	3.962	-0.65	-0.21	-0.46	-0.54	0.00	0.00	-0.37	-0.19	-0.56	-0.66	0.00	0.00	-0.65	-0.21	-0.46	-0.54	0.00	0.00

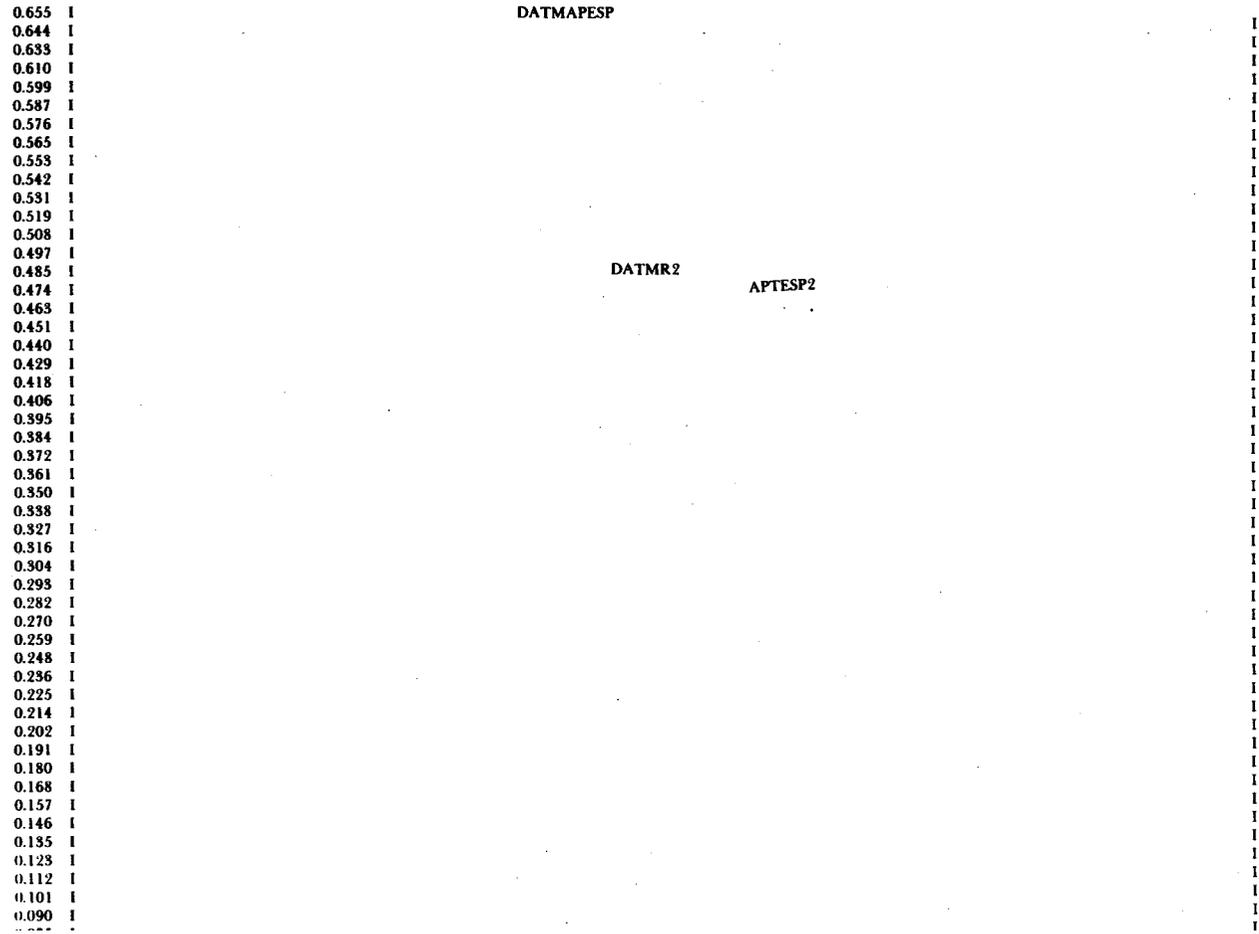
ELEMENTS SUPPLEMENTAIRES

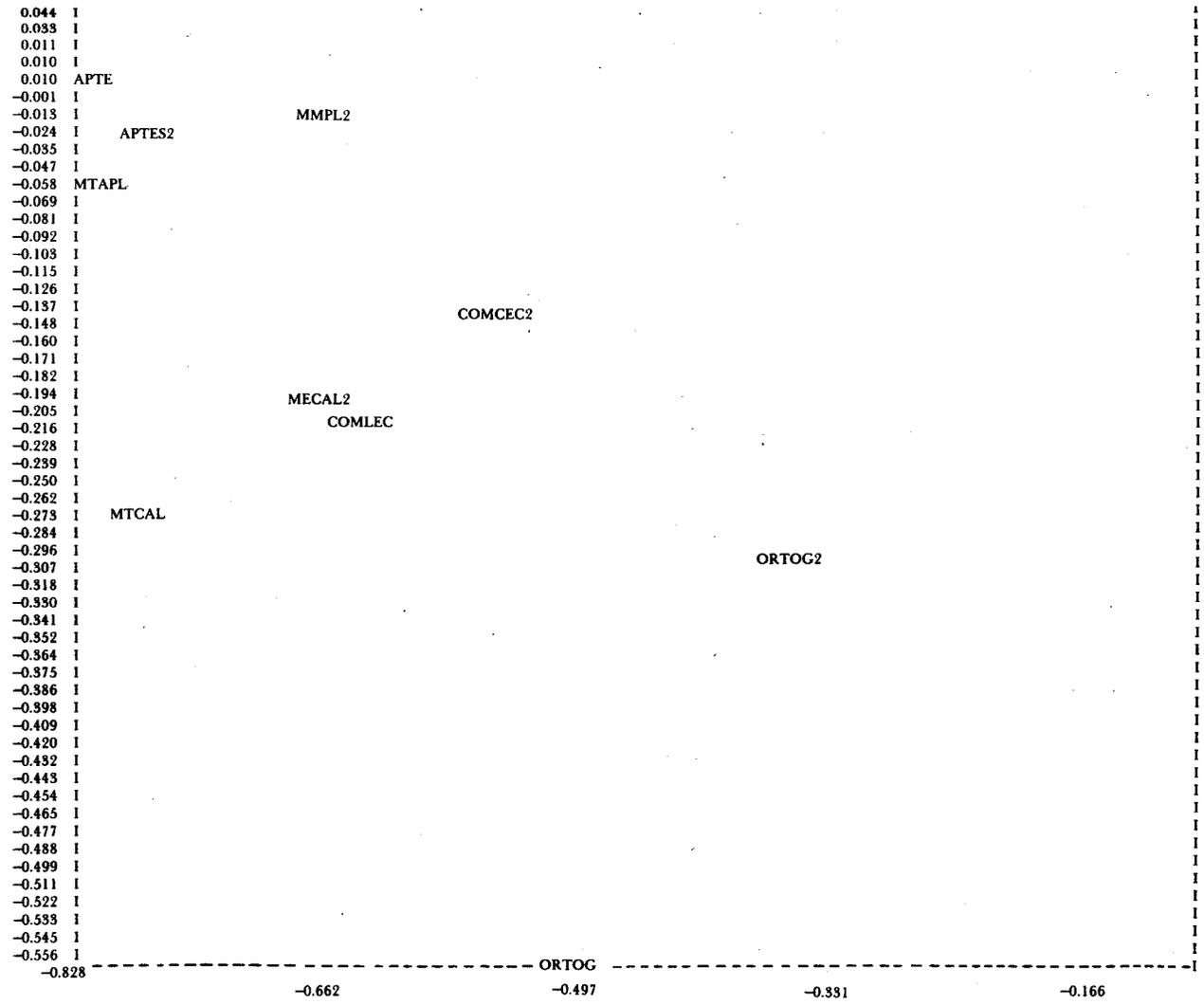
VALEUR-TEST POUR UNE CORRELATION SIGNIFICATIVE = .0227 (AU SEUIL 0.05)

COMLEC2	3.745	-0.55	-0.15	-0.21	-0.23	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-0.55	-0.15	-0.21	-0.23	0.00	0.00
ORTOG2	7.422	-0.36	-0.31	-0.12	0.13	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-0.36	-0.31	-0.12	0.13	0.00	0.00
MTCAL2	5.675	-0.66	-0.20	0.19	0.05	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-0.66	-0.20	0.19	0.05	0.00	0.00
MTAPL2	3.271	-0.67	0.02	0.17	-0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-0.67	-0.02	0.17	-0.02	0.00	0.00
APESP2	5.219	-0.40	0.48	-0.10	0.14	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-0.40	0.48	-0.10	0.14	0.00	0.00
APTES2	12.589	-0.76	-0.03	0.01	-0.06	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-0.76	-0.03	0.01	-0.06	0.00	0.00
DATMR2	7.844	-0.47	0.49	-0.02	-0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-0.47	0.49	-0.02	-0.02	0.00	0.00

FIGURA 2

PLAN DE PROJECTION DES 14 POINTS SUR LES AXES E ET 3
AXE 2/HORIZONTAL AXE 3/VERTICAL





BIBLIOGRAFIA

- Benzecri, J. P. «Sur le calcul des taux d'inertie dans l'analyse d'un questionnaire». *Les Cahiers de l'Analyse des Données*, 3, 1979, 377.
- Brown, B. W. and Saks, D. H. «The Production and Distribution of Cognitive Skills within Schools». *Journal of Political Economy*, 83 (3), 1975, 571-593.
- Chamberlain, G. «Multivariate Regression Models for Panel Data». *Journal of Econometrics*, 18, 1982, 5-46.
- Coleman, J. S.; Hoffer, Th. and Kilgore, S. *High School Achievement*. Nueva York, Basic Books, 1982 a.
- «Cognitive Outcomes in Public and Private Schools». *Sociology of Education*, 55, 1982 b, 65-76.
- Griliches, Z. «Estimating the Returns to Schooling: Some Econometric Problems». *Econometrica*, 45, 1977, 1-22.
- Hanushek, E. «Education and Race: An Analysis of the Educational Production Process». Heath, Lexington, Mass., 1972.
- Henderson, V.; Miezkowski, P. and Sauvageau, Y. «Peer-Group Effects and Educational Production Functions». *Journal of Public Economics*, 10, 1978, 97-106.
- Lalonde, R. J. «Evaluating the Econometric Evaluations of Training Programs with Experimental Data». *American Economic Review*, 76, 1986, 604-620.
- Modrego, A. y San Segundo, M. J. «Ecuaciones de rendimiento escolar para la evaluación de la Reforma de las Enseñanzas Medias». Abril, 1987.
- Murnane, R. J. «The Impact of School Resources on the Learning of Inner-City Children». Ballinger, Cambridge, Mass., 1975.
- San Segundo, M. «Empirical Studies of quality of schooling». Princeton, Ph. D. Thesis, 1985.
- Summers, A. A. and Wolfe, B. L. «Do Schools Make a Difference?» *American Economic Review* 67, (4), 1977, 639-652.