

# **El efecto de la inmigración en los resultados del sistema educativo. Una aproximación a partir de PISA 2003**

Javier Salinas Jiménez  
(Instituto de Estudios Fiscales)  
Daniel Santín González  
(Universidad Complutense de Madrid)

## **Resumen**

En este trabajo se analiza el grado de adaptación de los inmigrantes al sistema educativo español y el impacto de su llegada sobre los alumnos nativos. Para ello se estudia, utilizando la base de datos PISA 2003 de la OCDE, si el hecho de ser inmigrante influye en la probabilidad de repetición de curso de los propios inmigrantes y de los alumnos nativos. Este análisis se realiza utilizando una regresión logística ponderada para que los resultados sean interpretables a nivel de población. Posteriormente se estudia cuál es el efecto de la concentración de inmigrantes en las escuelas sobre el rendimiento académico de los alumnos en matemáticas, lectura y ciencias a partir de una regresión por mínimos cuadrados ponderados a nivel de alumno. Los resultados muestran como en general una mayor concentración de inmigrantes en la escuela supone una menor probabilidad de ser repetidor para los nativos. Además, el hecho de ser inmigrante aumenta considerablemente la probabilidad de ser repetidor.

## **1. Introducción**

Los resultados que aquí se presentan forman parte de un proyecto de investigación más amplio que tiene como objetivo analizar cómo la inmigración ha afectado en los últimos años a los presupuestos públicos tanto desde el lado de los ingresos como desde el lado de los gastos. Dentro de este marco general, se pretende analizar qué efectos ha provocado la llegada de la inmigración en el sistema educativo español. En esta ponencia nos centramos en el análisis de la repercusión que la entrada de alumnos inmigrantes ha tenido tanto en los resultados educativos de los alumnos españoles como en el de los propios inmigrantes. Para este análisis se utilizarán los datos españoles del Proyecto PISA (Programme for International Student Assessment) 2003 de la OCDE que es, tras la evaluación realizada en 2000, la segunda gran oleada de recogida de datos educativos, en esta ocasión especializada en la prueba de matemáticas.

Con esta base de datos, y teniendo en cuenta las peculiaridades de los datos que mencionaremos a continuación, en el trabajo se analizan dos aspectos clave en la producción educativa:

- En primer lugar, si la presencia de inmigrantes influye en la probabilidad de que un alumno, español o extranjero, repita curso, controlando por distintas variables socioeconómicas y escolares de los alumnos.
- En segundo lugar nos interesa analizar si los alumnos españoles que no han repetido curso ven disminuir su rendimiento por la presencia de aquellos inmigrantes que tampoco han repetido curso. Igualmente se analiza si los alumnos inmigrantes obtienen resultados académicos significativamente peores que los alumnos españoles por el hecho de ser inmigrantes<sup>1</sup>.

## **2. Datos**

El Proyecto PISA 2003 de la OCDE es, tras la evaluación realizada en 2000, la segunda gran oleada de recogida de datos educativos, en esta ocasión especializada en la prueba de matemáticas. PISA 2003 evaluó a 276.165 alumnos de 15 años de 41 países en las pruebas de matemáticas, ciencias y comprensión lectora y fue llevada a cabo

---

<sup>1</sup> Lo cual denotaría que a pesar de no haber repetido curso tienen dificultades adicionales respecto a un alumnado nativo con parecido nivel socioeconómico.

mayoritariamente en abril de 2003<sup>2</sup>. El proyecto PISA no evalúa las capacidades o destrezas como un conjunto de conocimientos del tipo “todo o nada”. Por el contrario, 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. Todos los alumnos<sup>3</sup> completaron cuestionarios acerca de su entorno socioeconómico y sus experiencias dentro y fuera del ámbito escolar. De igual manera, los directores del colegio rellenaron cuestionarios sobre las prácticas educativas que llevaban a cabo.

En España se evaluaron a 10.761 alumnos repartidos en 383 centros escolares. Además, las Comunidades Autónomas de Castilla y León, Cataluña y País Vasco han participado en esta oleada con una muestra ampliada. Ello permite que, además de poder obtener resultados para toda España, podamos identificar efectos individuales para estas Comunidades.

El muestreo se realizó en dos etapas (OCDE 2005a, p. 124): en primer lugar se seleccionaron las escuelas que formarían parte de la muestra. En segundo lugar se seleccionaron aleatoriamente 35 alumnos de cada escuela. La población objetivo, formada concretamente por el alumnado comprendido entre las edades de 15 años y tres meses y 16 años y dos meses (OCDE 2005a, p.124), puede estar comprendida en diferentes cursos. En el caso de España la distribución es la siguiente:

**Tabla 1: Distribución del alumnado en PISA 2003.**

	Curso					Total
	1º ESO	2º ESO	3º ESO	4º ESO	1º Bachillerato	
Nº de alumnos	4	259	2576	7951	1	10791
Porcentaje	0,0%	2,4%	23,9%	73,7%	0,0%	100%

Elaboración propia a partir de PISA 2003.

<sup>2</sup> Para un análisis descriptivo de los principales resultados del Informe PISA 2003 puede acudir a OECD (2004).

<sup>3</sup> En concreto, según OECD (2005a, p. 20), el Informe PISA demanda una tasa de participación superior al 80% del alumnado seleccionado de 15 años para que la escuela sea incluida en el estudio.

### ***Repetidores y alumnos en cursos retrasados***

Cabe destacar que, para España, todos los alumnos de la muestra PISA 2003 nacieron en el año 1987. Los alumnos de más de 15 años y 3 meses que no estaban en 4º ESO eran repetidores o iban retrasados con respecto al curso que les correspondería. Para no asistir a 4º ESO pueden existir dos razones:

- Que el alumno sea repetidor.
- Que el alumno no sea repetidor pero por algún motivo haya sido escolarizado tardíamente en un curso inferior.

Las respuestas de los alumnos en PISA 2003<sup>4</sup> permitieron identificar si existía una correspondencia entre aquéllos que declaraban haber repetido y el curso en el que estaban, para de esta forma comprobar que la hipótesis del retraso era cierta. Con esta finalidad distinguimos entre alumnos no repetidores y alumnos repetidores (aquéllos que han repetido curso en una o más ocasiones desde primero de Educación Primaria hasta 4º de ESO). Los resultados se ofrecen en la tabla 2:

***Tabla 2: Alumnos no repetidores y repetidores***

	No repetidores	Repetidores	Datos Perdidos	Total
1º ESO	3	1	0	4
2º ESO	9	243	7	259
3º ESO	103	2432	41	2576
4º ESO	7785	36	130	7951
1º BACHILL	1	0	0	1
TOTAL	7901	2712	178	10791

Puede observarse que la gran mayoría de los alumnos que declaran haber repetido van un curso retrasados. Por otro lado, podemos observar 115 alumnos que declaran no haber repetido y sin embargo asisten a cursos retrasados respecto a su año de nacimiento.

<sup>4</sup> A partir de la pregunta del cuestionario del alumno Q22: “¿Has repetido alguna vez un curso académico?” pueden identificarse estas variables:

- ISCED 1 (Educación Primaria): Las respuestas posibles eran; no nunca, sí una vez, sí dos o más veces.
- ISCED 2 (Desde 1º ESO a 4º ESO): Las respuestas posibles eran; no nunca, sí una vez, sí dos o más veces.

(ISCED son las siglas que la OCDE aplica al Internacional Standard Classification of Education que se utiliza para realizar comparaciones internacionales de sistemas educativos).

Como se decía anteriormente la explicación fundamental para la situación de estos alumnos es que hayan sido escolarizados de forma tardía. Se pueden observar también 36 alumnos que declaran haber repetido y que sin embargo asisten al curso correcto. La razón para este porcentaje de alumnos es justo la contraria al caso anterior, es decir, alumnos que fueron escolarizados un año antes al que les correspondía y posteriormente repitieron curso<sup>5</sup>.

Para analizar si el hecho de ser inmigrante está asociado a ir retrasado en la escuela, habiendo repetido curso o no, utilizamos la información disponible para España en la base de datos PISA 2003 de alumnos de 15 años. Esta información permite comparar los resultados del alumnado nativo con las distintas definiciones de alumnado inmigrante o extranjero posibles; en concreto, hemos tomado las categorías que por el origen de nacimiento estableció el informe PISA 2000, añadiendo además una nueva categoría:

**Nativo:** Si el alumno y al menos uno de sus padres nació en España. También se considera nativo al alumno nacido en el extranjero cuando al menos uno de sus padres nació en España.

**No-Nativo:** Cuando ni el alumno ni al menos uno de sus padres nació en España.

**Primera generación:** Si el alumno nació en España pero sus padres nacieron fuera de España.

**Indeterminado:** Cuando no existe información suficiente para clasificar al alumno en alguna de las categorías anteriores. Es fácil sospechar que cuando un alumno a los 15 años no declara su origen o el de alguno de sus padres es porque tiene algún tipo de prevención a declarar un origen distinto al español. Consideraremos, por tanto, que la mayor parte de este alumnado podría ser también de origen inmigrante. Como se puede observar en los resultados que siguen (tabla 3) el alumno clasificado en esta categoría obtiene unos resultados, en la mayoría de casos significativamente distintos del alumnado nativo.

A partir de estas definiciones se puede comprobar cuál es la distribución por curso del alumnado atendiendo a su procedencia. Los resultados se muestran en la tabla 3.

---

<sup>5</sup> En ambos casos otra razón posible es que mintieran en la respuesta.

**Tabla 3: Origen del alumno y grado al que asisten**

**Tabla de contingencia Country of birth \* Grade Q1a**

		Grade Q1a					Total
		1º ESO	2º ESO	3º ESO	4º ESO	1º BACH	
Native students	Recuento	3	234	2409	7713	0	10359
	% de Country of birth	,0%	2,3%	23,3%	74,5%	,0%	100,0%
	% de Grade Q1a	75,0%	90,3%	93,5%	97,0%	,0%	96,0%
	% del total	,0%	2,2%	22,3%	71,5%	,0%	96,0%
First-Generation students	Recuento	0	2	15	31	0	48
	% de Country of birth	,0%	4,2%	31,3%	64,6%	,0%	100,0%
	% de Grade Q1a	,0%	,8%	,6%	,4%	,0%	,4%
	% del total	,0%	,0%	,1%	,3%	,0%	,4%
Non-native students	Recuento	1	16	99	108	1	225
	% de Country of birth	,4%	7,1%	44,0%	48,0%	,4%	100,0%
	% de Grade Q1a	25,0%	6,2%	3,8%	1,4%	100,0%	2,1%
	% del total	,0%	,1%	,9%	1,0%	,0%	2,1%
Incierto	Recuento	0	7	53	99	0	159
	% de Country of birth	,0%	4,4%	33,3%	62,3%	,0%	100,0%
	% de Grade Q1a	,0%	2,7%	2,1%	1,2%	,0%	1,5%
	% del total	,0%	,1%	,5%	,9%	,0%	1,5%
Total	Recuento	4	259	2576	7951	1	10791

Puede observarse como los alumnos nativos asisten en mayor medida al curso que les corresponde mientras que no ocurre lo mismo con los alumnos inmigrantes.

Podemos también preguntarnos si la distribución del alumnado que va retrasado y declara que es repetidor es la misma entre los inmigrantes y los nativos. Para ello hemos realizado un análisis de tablas de contingencia. La tabla 4 muestra los resultados.

**Tabla 4: Distribución de los alumnos que declaran no haber repetido y van un curso retrasado.**

	NATIVO	PRIM_GEN	NO NATIVO	INCIERTO	TOTAL
NO REPETIDOR RETRASADO	65*	2	41**	7**	115
ALUMNO EN SU CURSO	10294**	46	184*	152*	10676
	0,63%	4,17%	18,22%	4,40%	10791

\* La proporción de la celda es significativamente inferior a la esperada.

\*\* La proporción de la celda es significativamente superior a la esperada.

Puede comprobarse que más del 18% de los inmigrantes no nativos están escolarizados en un curso inferior al que le correspondería por su edad. Los alumnos de primera generación son los únicos que no tienen un resultado significativamente distinto al de la frecuencia esperada.

Dado este comportamiento y la distribución del alumnado en PISA 2003, hemos creado, en primer lugar, una variable RETRASO que toma valor 1 cuando el alumno es repetidor o va retrasado y valor 0 cuando, por edad, va en el curso que le corresponde. De esta forma podemos analizar, controlando por el efecto de distintas variables escolares e individuales, si el tamaño del grupo en la escuela, el porcentaje de inmigrantes en la escuela y el hecho de ser inmigrante influye en la probabilidad de que un alumno repita curso. Para realizar este análisis hemos utilizado una regresión logística a nivel de alumno ponderada por los pesos individuales para que sea representativa a nivel de toda la población.

En un análisis posterior, comprobamos si los resultados en matemáticas, lectura y ciencias de los alumnos que van bien por curso resultan afectados porque en su escuela haya mayor o menor proporción de inmigrantes.

### **3. Resultados académicos y variables explicativas**

#### ***3.1.- Resultados académicos y valores plausibles***

El verdadero producto de la educación recibida por un individuo es muy difícil de medir empíricamente debido a su intangibilidad inherente. La educación no se compone únicamente de la capacidad de repetir información y resolver cuestiones sino que comprende además la capacidad de razonar, el comportamiento y los buenos modales en la sociedad y el pensamiento crítico, todos ellos de muy difícil cuantificación. A pesar de la naturaleza multi-producto de la educación, la mayoría de los trabajos han utilizado los resultados de rendimiento en pruebas objetivas por distintas razones. Entre ellas cabe destacar que son difíciles de falsificar y responden a las exigencias administrativas para pasar curso e incluso acceder a la universidad. Pero quizás, siguiendo a Hoxby (1999), lo más importante sea que tanto políticos como padres de alumnos usan este

criterio para evaluar el producto educativo y utilizan la información que se deriva para elegir escuela, tomar decisiones de inversión e incluso lugar de residencia.

En este contexto de evaluación del producto educativo, existe un acuerdo prácticamente unánime en todos los países del mundo en considerar que disciplinas como las matemáticas, las ciencias o la capacidad lectora, son una parte importante del currículo escolar y se consideran asignaturas fundamentales para la educación de los alumnos. Esto es así porque estas materias son un pilar básico para la futura integración del individuo en el mercado de trabajo y le preparan para afrontar con éxito el reto tecnológico, científico y de la sociedad de la información que plantea el mercado de trabajo actual. En este trabajo utilizamos como indicadores del rendimiento de cada alumno los resultados en matemáticas, comprensión lectora y ciencias en la prueba objetiva de PISA 2003<sup>6</sup>.

Hay que tener en cuenta que los informes internacionales como PISA 2003 utilizan el concepto de valores plausibles para medir el rendimiento de los alumnos. Al contrario que en otras ciencias como la física o la economía, las medidas del conocimiento educativo tienen un amplio margen de error debido a que el concepto que se quiere medir es abstracto y además la medida está sujeta a las especiales circunstancias del alumno y de su entorno el día del examen. Todo ello supone también que, en educación, el error de medida no es independiente de la posición del alumno en la distribución de resultados. Los alumnos con muy malos o muy buenos resultados tienen mayores errores de medida y mayor asimetría en su distribución que los alumnos con resultados medios. Es por ello que en la teoría de respuesta a los ítems de medición del nivel educativo, en vez de trabajar con un valor medio se trabaja con valores plausibles, entendidos éstos como una representación del rango de habilidades que tiene cada estudiante. Los valores plausibles son valores aleatorios extraídos de la función de distribución de resultados estimada a partir de las contestaciones en cada prueba<sup>7</sup>. PISA 2003 proporciona cinco valores plausibles para cada alumno.

---

<sup>6</sup> La escala continua utilizada para medir el nivel de conocimientos de los alumnos en PISA 2003 tiene un significado educativo concreto. Así, por ejemplo, para el resultado en matemáticas existen 6 niveles de comprensión en función de los conocimientos y capacidades el alumno tiene en cada uno de ellos.

<sup>7</sup> Para una revisión de la literatura de los valores plausibles puede acudir a Mislevy (1991) y Mislevy *et al.* (1992). También puede acudir a OECD (2005a, capítulo 5).

El análisis con valores plausibles supone realizar el análisis cinco veces y promediar el resultado obtenido. Tal y como se afirma en OCDE (2005a, p. 79): “El análisis de los datos (procedentes de PISA 2003) nunca debe partir de promediar los cinco valores plausibles a nivel de alumno y a continuación calcular el estadístico de interés”.

### 3.2.- Variables explicativas

- *Estatus Socioeconómico del alumno (ESCS)*: Esta variable fue construida por los analistas del proyecto PISA agrupando las respuestas a distintas preguntas planteadas en el cuestionario de la evaluación<sup>8</sup>. Así, la variable *Economic, Social and Cultural Status* (ESCS) es un índice obtenido a partir de tres variables relacionadas con el contexto socioeconómico familiar. La primera variable es el nivel educativo más alto de cualquiera de los padres medida a partir del *International Standard Classification of Education* (ISCED, OECD, 1999). La segunda variable es el índice más alto de ocupación laboral de cualquiera de los dos padres medido a partir de la *Internacional Socio-economic index of Occupational Status* (ISEI, Ganzeboom *et al.*, 1992). La tercera variable es un índice de posesiones educativas relacionadas con la economía del hogar<sup>9</sup>. Dado que los valores de esta variable eran positivos y negativos fue reescalada sumándole el valor mínimo más uno. Valores más positivos indican mayor estatus socioeconómico.

- *Efecto Compañeros (EFCO)*: La idea que subyace en este concepto es que el nivel de conocimientos que un alumno asimila depende directamente de las características de sus compañeros<sup>10</sup>. Este hecho supone que un alumno de alto rendimiento que asiste a una escuela deja de generar una externalidad positiva sobre los alumnos de otra escuela. Sin embargo, la asignación óptima de alumnos a lo largo de las escuelas, así como que

---

<sup>8</sup> No es el objetivo de este trabajo realizar una extensa descripción de las variables que recoge el proyecto PISA ni de cómo esta información fue agrupada. En este trabajo tan solo se presenta una breve descripción para una adecuada comprensión del análisis. No obstante, para el estudio de todas las variables, los métodos de muestreo empleados y toda la información de cómo fue desarrollado el proceso estadístico de agrupación de variables puede acudir a OECD (2005a, 2005b).

<sup>9</sup> Este último índice se construyó a partir de las respuestas que el alumno dio sobre si poseía o no en casa los siguientes elementos: lugar de estudio, habitación propia, lugar de estudio tranquilo, ordenador para tareas escolares, software educativo, conexión a internet, calculadora propia, libros de literatura clásica, libros de poesía, trabajos de arte, libros de ayuda educativa, diccionario, lavavajillas y consola de videojuegos. Además, los alumnos españoles fueron preguntados para construir este índice acerca de si poseían video, DVD y video consola (Play Station, X-Box, Nintendo o similar).

<sup>10</sup> Para una revisión véase, Betts (2000), Hanushek *et al.* (2001).

padres y alumnos acepten esa asignación, es una cuestión que no resuelve el mercado y que no se incorpora claramente en los modelos de asignación de recursos ni está resuelta por la teoría en economía de la educación. El objetivo social educativo puede alcanzarse en menor o mayor medida en función del grado de heterogeneidad que exista dentro del aula. Sin embargo, los alumnos mejor dotados tanto en habilidad como en recursos económicos tenderán a segregarse de los peor dotados<sup>11</sup> que a su vez tratarán de buscar las externalidades positivas que les proporcionan los alumnos de este primer grupo. Esta variable se ha cuantificado como el nivel medio en la variable ESCS de los compañeros que comparten la misma escuela que el alumno evaluado. Valores más positivos indican un mejor efecto compañeros.

- *Calidad de los Recursos Educativos (SMATEDU)*: Es también un índice derivado a partir de las respuestas del director de la escuela a siete preguntas relacionadas con el grado de disponibilidad de determinados recursos. Más concretamente se preguntó en qué medida el colegio tenía necesidad de recursos en los siguientes elementos: materiales educativos, ordenadores para usos didácticos, software educativo, calculadoras, libros y material de biblioteca, recursos audiovisuales y material y equipo de laboratorio. Valores más positivos indican mejor calidad de los recursos escolares.

- *Disciplina en el aula*

◆ *DISCLIM*: Este índice se construyó a partir de las respuestas que cada alumno declara sobre la frecuencia con la que ocurren en clase los siguientes hechos: a) los alumnos no escuchan lo que el profesor dice, b) hay ruido y desorden, c) el profesor tiene que esperar mucho tiempo hasta que los alumnos se están quietos d) los alumnos no pueden trabajar bien e) se tarda un largo período desde que empieza la clase hasta que los alumnos comienzan a trabajar. El índice fue construido de tal forma que valores positivos de la variable indican la percepción individual del alumno de un clima de trabajo positivo. De esta manera se pretende contrastar en qué medida el entorno que proporciona el resto de compañeros y el profesor en el

---

<sup>11</sup> Este fenómeno es conocido en Economía de la Educación como “retirar la nata” (del término inglés Cream Skimming Effect). No está claro en los resultados empíricos si los efectos positivos que recibe un buen alumno que se cambia a una escuela con mejores compañeros compensa la pérdida de bienestar que sufren los alumnos que se quedan sin el compañero bueno. Los trabajos científicos que abordan el tema tanto de los “cheques escolares” como del “efecto compañeros” y por tanto ligados a los efectos que origina la movilidad de los alumnos son numerosos. Un buen trabajo de referencia es el de Dills (2005).

aula ayuda, perjudica o es indiferente en el rendimiento del alumno y en la probabilidad de que repita curso.

- ◆ *LLEGATARDE*: Al alumno se le pregunta en la cuestión 28: “En las últimas dos semanas que asististe al colegio ¿Cuántas veces llegaste tarde al colegio?”. Las posibles respuestas eran: nunca; una o dos veces; tres o cuatro veces; cinco o más veces. Esta variable fue codificada de forma ordinal asignando valor 1 a los que nunca llegaron tarde y 4 a los que lo hicieron más de cinco veces.

- *Posesiones en el Hogar*

- ◆ *ORDENAD\_CASA*: Mide las posesiones informáticas en el hogar. Recoge el índice *Computer facilities at home*. Este índice se deriva a partir de las respuestas de los alumnos a: ¿Cuál de los siguientes elementos tienes en casa? Las posibles respuestas eran: un ordenador para realizar deberes; software educativo; conexión a internet.
- ◆ *EDUC\_CASA*: Mide las posesiones educativas en el hogar. Recoge el índice *Home Possessions*. Este índice se deriva a partir de las respuestas de los alumnos a: ¿Cuál de los siguientes elementos tienes en casa? Las posibles respuestas eran: Escritorio para estudiar; un lugar tranquilo para estudiar; tu propia calculadora; libros para ayudarte con los deberes; un diccionario.
- ◆ *CULT\_CASA*: Mide las posesiones culturales en el hogar. Recoge el índice *Cultural Possessions*. Este índice se deriva a partir de las respuestas de los alumnos a: ¿cuál de los siguientes elementos tienes en casa? Las posibles respuestas eran: literatura clásica; libros de poesía; trabajos de arte.

Dado que los valores de estas variables presentaban valores positivos y negativos todas las variables fueron reescaladas para que presentaran valores positivos en aras de una interpretación más adecuada.

- ◆ *CONSOLA*: En la cuestión 17 se pregunta a los alumnos si poseen determinados elementos. Esta variable toma valor 1 si el alumno declara poseer consola de videojuegos (Playstation, X-Box, Nintendo, etc.) y 0 si no la posee.
- ◆ *PREESCOLAR*: La variable se construye a partir de la pregunta 20 “¿asististe a preescolar?”. Tras un análisis exploratorio las respuestas fueron recodificadas en dos categorías: 1 – si el alumno asistió a preescolar un año o más; 0 – si el alumno no asistió a preescolar o lo hizo por menos de un año.

- *Tipo de familia (MONOPARENTAL)*: PISA 2003 clasifica a las familias como nucleares, mixtas, monoparentales y otras. Esta variable toma valor 1 si el alumno pertenece a una familia monoparental (cuando el alumno vive con sólo uno de los siguientes: su padre, su madre, un padrastro, una madrastra) o a la categoría “otro tipo de familias” y 0 si el alumno pertenece a una familia nuclear o mixta.

- *Tamaño del Grupo (TAMGRUPO)*: Se calcula a partir de la pregunta 36 del cuestionario de los alumnos: “En media, ¿cuántos alumnos asisten a tu clase de matemáticas?” Las respuestas de todos los estudiantes de la misma escuela son promediadas para tener una aproximación al tamaño de los grupos en esa escuela. En diferentes estudios esta variable se ha introducido en el análisis como un input entendiendo que grupos más reducidos supondrían mejoras en el rendimiento. Sin embargo, en economía de la educación no se conoce cuál es el valor óptimo para llevar a cabo la producción educativa. Esto es, no existe ningún resultado científico definitivo que señale el número de alumnos máximo que puede recibir educación en un aula sin que se deteriore el aprendizaje de los demás. Tampoco existen resultados definitivos, sino más bien una elevada controversia, acerca del efecto que el tamaño de la clase tiene en el aprendizaje de los alumnos [Hanushek, 1986, 1997, 2003; Krueger, 1999; Hoxby, 2000, Rivkin *et al.* 2005]. Dado que no existe una relación teórica del efecto directo de esta variable en los resultados hemos optado por incluirla en el análisis utilizando además del tamaño de la clase el cuadrado de esta variable [TAMGRUPO2].

- *Inmigrante*: Distingue cuatro categorías por el lugar de nacimiento del alumno. La categoría de referencia es ser nativo, esto es, alumnos nacidos en el país en el que recibe la educación y con al menos un padre también nacido en el mismo país. También es considerado nativo el alumno nacido en el extranjero con al menos un padre nacido en el país en el que recibe su educación.

- ◆ PRIM\_GENERACION: Toma valor 1 para los alumnos nacidos en el país en el que reciben la educación pero con ambos padres nacidos en el extranjero y valor 0 para el resto de categorías.
- ◆ NO NATIVO: Toma valor 1 cuando tanto el alumno como los padres nacen en un país distinto al del que el alumno se encuentra escolarizado y valor 0 para el resto de categorías.

- ◆ INCIERTO: Toma valor 1 para aquellos alumnos que no se pueden clasificar en alguna de las categorías anteriores y 0 para el resto de categorías.
- ◆ PORC\_INMIG: Es el porcentaje de alumnos inmigrantes (alumnos que no pertenecen a la categoría “nativos”) encuestados respecto al total de alumnos encuestados en la escuela. A pesar de que sólo se tiene información del porcentaje que representan los inmigrantes de 15 años, esta variable es una *proxi* del porcentaje de inmigrantes en la escuela. Se pretende comprobar si un alumno resulta perjudicado por ser educado junto a alumnos inmigrantes. Se calculó también la variable PORC\_INMIG2 como el cuadrado de la variable anterior y PORC\_INMIG\_PUB que recoge el efecto de si el porcentaje de inmigrantes en la escuela pública tiene un efecto igual o no a este mismo porcentaje en la privada.

- *Titularidad de la Escuela*: Es interesante analizar si la producción educativa por parte de una escuela privada, privada concertada (financiación pública y titularidad privada) o pública influye en la eficiencia de los alumnos considerados. La categoría de referencia es PRIVADA: Escuela de titularidad y financiación privada.

- ◆ PUBLICA: Toma valor 1 si el alumno asiste a una escuela pública y 0 para el resto de categorías.
- ◆ CONCERTADA: Toma valor 1 si el alumno asiste a una escuela concertada y 0 para el resto de categorías.

- *Sexo (SEXO)*: Toma valor 1 para el sexo masculino y 0 para el femenino.

- *Comunidad Autónoma a la que pertenece el alumno*: Se tiene en cuenta si por el hecho de pertenecer a una CC. AA. los resultados varían respecto a los de otra. En ello podría incidir una política educativa distinta. La categoría de referencia es “resto de España”.

- ◆ CASTLEON: Toma valor 1 para los alumnos que asisten a una escuela en Castilla y León y 0 para otras CC. AA.
- ◆ CATALUÑA: Toma valor 1 para los alumnos que asisten a una escuela en Cataluña y 0 para otras CC. AA.
- ◆ PVASCO: Toma valor 1 para los alumnos que asisten a una escuela en el País Vasco y 0 para otras CC. AA.

- *Tamaño del municipio*: Se analiza también si las escuelas radicadas en municipios grandes, medianos o pequeños tienen mejores resultados. La categoría de referencia es Escuela en municipio de menos de 15.000 habitantes.

- ◆ CIUDAD: Toma valor 1 para los alumnos que reciben su educación en municipios de entre 15.000 y 100.000 habitantes y 0 para el resto de categorías.
- ◆ GRANCIUDAD: Toma valor 1 para los alumnos que reciben su educación en municipios de más de 100.000 habitantes.

#### **4. Modelo de regresión logística para explicar la repetición de curso.**

La regresión logística es un método estadístico que generaliza los modelos lineales generales y que ha sido y es frecuentemente utilizado para modelar la probabilidad de que ocurra un suceso para una variable de respuesta binaria. En nuestro caso la variable categórica es la RETRASO que recibe valor 1 si el alumno va retrasado respecto al curso que le correspondería y 0 si no va retrasado.

En un modelo de regresión logística la relación entre las variables explicativas  $X$  y la probabilidad asignada a que el evento de interés suceda se denomina función logística. Para este propósito utilizaremos la función logit:

$$p = \frac{\exp(\beta X)}{1 + \exp(\beta X)}$$

Los parámetros  $\beta$  de este modelo se estiman maximizando la función de verosimilitud correspondiente, teniendo en cuenta la ponderación asignada a cada alumno, mediante un algoritmo iterativo tipo Newton-Raphson.

La tabla 5 muestra los resultados de la regresión logística ponderada para explicar el hecho de que un alumno vaya retrasado.

**Tabla 5: Probabilidad de ser un alumno retrasado o repetidor para todos los alumnos examinados en PISA 2003 en toda España.**

Variable	Valor	Desviación típica	Chi-cuadrado de Wald	Pr > Chi²	Valores odds
Intersección	1,328	0,132	101,437	< 0,0001	3,7721
Llegatarde	0,353	0,005	5890,945	< 0,0001	1,4227
escspos	0,785	0,034	519,597	< 0,0001	2,1914
escspos2	-0,133	0,004	1083,196	< 0,0001	0,8755
escspos_mean	<b>-0,035</b>	0,012	9,340	<b>0,002</b>	0,9653
orded_casa	-0,222	0,005	2210,453	< 0,0001	0,8008
cultura_casa	-0,108	0,005	533,420	< 0,0001	0,8979
educ_home	-0,179	0,006	768,429	< 0,0001	0,8358
disclim	-0,439	0,005	8205,163	< 0,0001	0,6445
PORC_INMIG	<b>-0,011</b>	0,002	27,089	<b>&lt; 0,0001</b>	0,9889
PORC_INMIG_PUB	<b>0,007</b>	0,002	15,067	<b>0,000</b>	1,0072
porc_inmig_2	0,000	0,000	0,531	0,466	1,0000
TAMAÑO_GRUPO	-0,022	0,009	6,006	0,014	0,9781
TAMG2	0,000	0,000	0,022	0,882	1,0000
SCMATEDU	-0,050	0,005	95,956	< 0,0001	0,9514
castleon	0,447	0,018	589,201	< 0,0001	1,5628
cataluña	-1,658	0,018	8546,849	< 0,0001	0,1905
pvasco	-0,233	0,023	106,202	< 0,0001	0,7919
sexo	0,406	0,009	1935,814	< 0,0001	1,5001
publica	0,225	0,022	101,864	< 0,0001	1,2519
concertada	-0,174	0,021	67,930	< 0,0001	0,8400
ciudad	0,023	0,012	3,758	0,053	1,0231
gran_ciudad	0,118	0,012	105,731	< 0,0001	1,1255
PRIM_GENERACION	<b>0,452</b>	0,059	57,858	<b>&lt; 0,0001</b>	<b>1,5713</b>
NO_NATIVO	<b>0,792</b>	0,027	846,821	<b>&lt; 0,0001</b>	<b>2,2072</b>
INCIERTO	<b>1,003</b>	0,043	549,487	<b>&lt; 0,0001</b>	<b>2,7256</b>
MONOPARENTAL	0,337	0,012	799,849	< 0,0001	1,4002
CONSOLA	0,271	0,011	653,321	< 0,0001	1,3113
PRESCHOOL	-0,167	0,012	191,252	< 0,0001	0,8460

Los resultados de la tabla 5 muestran un número de resultados interesantes.

- ◆ El hecho de ser inmigrante en cualquiera de sus tres variantes aumenta el riesgo de ser repetidor de curso respecto a los nativos en un 57% para los alumnos de primera generación, en un 120% para los alumnos no nativos y en un 172% para el alumnado de origen incierto.
- ◆ No se encuentra una relación clara entre el tamaño del grupo y la probabilidad de ser repetidor ya que el efecto significativo encontrado es que cuanto mayor es el tamaño del grupo menor es la posibilidad de repetir.
- ◆ Un mayor porcentaje de inmigrantes en la escuela supone una menor probabilidad de repetir curso para cualquier alumno si bien este efecto aunque se mantiene

disminuye en la escuela pública. Este resultado está claramente en relación con el ya comentado de que los inmigrantes tienen más probabilidad de ser repetidores.

- ◆ Un mejor nivel socioeconómico del alumno y de los compañeros de escuela influye a favor de ser no repetidor.
- ◆ Para un alumno español varón residente en una gran ciudad del resto de España que asiste a una escuela pública, vive en una familia no monoparental, asistió a preescolar, no tiene consola de videojuegos, no llega tarde a la escuela y que ostenta valores medios en todas las variables cuantitativas, la probabilidad de ser repetidor es del 28,9%. En el caso de un alumno no nativo con las mismas características la probabilidad aumenta al 47,3% mientras que si se trata de un alumno de primera generación esta probabilidad sería del 39,0%.
- ◆ Para un alumno español varón residente en una gran ciudad del resto de España que asiste a una escuela pública, vive en una familia monoparental, no asistió a preescolar, tiene consola de videojuegos, llegó tres o cuatro veces tarde a la escuela en el último mes y que ostenta valores correspondientes al primer cuartil en todas las variables socioeconómicas, la probabilidad de ser repetidor es del 80,0%. En el caso de un alumno no nativo con las mismas características la probabilidad aumenta al 89,8% mientras que si se trata de un alumno de primera generación esta probabilidad baja al 86,2%.
- ◆ Si lleváramos al alumno anterior a una escuela concertada con una distribución del efecto compañeros y de los recursos escolares en el tercer cuartil, el alumno asistiera puntualmente a la escuela y no jugara a la consola de videojuegos su probabilidad de ser repetidor bajaría al 37,7% para el alumno español, al 57,2% en el caso del no nativo y al 48,7% en el caso del alumno de primera generación. Se puede comprobar por tanto como, siendo significativo ser inmigrante para tener más probabilidades de ser repetidor, esta probabilidad puede compensarse modificando otras variables tanto para los alumnos españoles como inmigrantes.
- ◆ El modelo tiene un grado de ajuste elevado; 77,76%, tal y como muestra la matriz de confusión obtenida y la curva ROC con un área bajo la curva igual a 0,782.

## 5. Modelo explicativo del rendimiento académico entre los alumnos no repetidores.

Una vez que hemos analizado el hecho de ir retrasado cabe ahora explicar los factores que afectan al rendimiento de los alumnos que van en el curso que les corresponde y si entre esos factores se encuentra asistir a una escuela con mayor porcentaje de inmigrantes. Para este fin utilizaremos mínimos cuadrados ponderados a nivel de alumno para explicar el rendimiento en matemáticas, lectura y ciencias. Dado que el rendimiento se mide en valores plausibles, cada análisis debe realizarse cinco veces obteniendo el valor final y haciendo la media de los coeficientes estimados. En la tabla 6 se muestran, por razones de simplicidad, sólo las estimaciones finales.

**Tabla 6: Rendimiento académico entre los alumnos no repetidores**

VARIABLES	Matem	P-Valor	Lectura	P-Valor	Ciencias	P-Valor
(Constante)	171,1910	0,0000	186,4799	0,0000	137,0592	0,0000
Llegatarde	-7,8268	0,0000	-8,4223	0,0000	-9,4000	0,0000
escspos	7,2846	0,3272	22,3624	0,0097	16,7572	0,0649
escspos2	0,0201	0,6549	-1,8283	0,0575	-0,7333	0,4718
escspos_mean	<b>32,7452</b>	<b>0,0000</b>	<b>27,3352</b>	<b>0,0000</b>	<b>29,1724</b>	<b>0,0000</b>
orded_casa	4,1189	0,0000	4,7860	0,0000	3,4599	0,0045
cultura_casa	5,1629	0,0000	5,0972	0,0000	6,1078	0,0000
educ_home	2,3327	0,1913	3,9453	0,0387	4,1628	0,1603
disclim	8,3253	0,0000	8,7650	0,0000	9,7582	0,0000
PORC_INMIG	0,2409	0,3920	-0,3487	0,4883	-0,3349	0,4684
PORC_INMIG_PUB	-0,2305	0,4533	-0,0749	0,5996	0,2545	0,5314
porc_inmig_2	-0,0115	0,1891	-0,0001	0,6940	-0,0044	0,6535
TAMAÑO_GRUPO	5,8632	0,0009	4,4978	0,0432	7,8296	0,0001
TAMG2	-0,1235	0,0023	-0,0726	0,1481	-0,1618	0,0003
SCMATEDU	3,8528	0,0004	4,4233	0,0001	5,4909	0,0000
castleon	14,7700	0,0001	14,8353	0,0011	8,9870	0,0577
cataluña	-11,2540	0,0000	-19,3628	0,0000	-9,7339	0,0008
pvasco	8,1483	0,0376	9,0713	0,0410	-13,7258	0,0031
sexo	24,3821	0,0000	-24,5477	0,0000	18,4880	0,0000
escuela publica	14,5302	0,0001	10,2696	0,0258	11,6909	0,0126
escuela concertada	2,8618	0,3676	0,4212	0,4312	0,2688	0,5285
escuela en ciudad	-6,7451	0,0071	-5,9477	0,0161	-5,4280	0,0384
gran_ciudad	-0,2055	0,8233	2,1296	0,3747	4,3519	0,1287
PRIM_GEN	<b>-32,7278</b>	<b>0,0154</b>	<b>-25,9990</b>	<b>0,0935</b>	<b>-39,6259</b>	<b>0,0234</b>
no nativo	<b>-32,4376</b>	<b>0,0000</b>	<b>-26,5138</b>	<b>0,0003</b>	<b>-31,8644</b>	<b>0,0001</b>
Origen incierto	-18,5802	0,1141	<b>-47,5594</b>	<b>0,0004</b>	<b>-30,5135</b>	<b>0,0510</b>
MONOP	-4,0226	0,1555	-4,2228	0,1712	1,9332	0,5005
CONSOLA	-5,9049	0,0064	-3,4494	0,1529	-4,6158	0,0958
Preescolar	14,1542	0,0000	8,8296	0,0090	1,1432	0,5799
R2 corregida	0,18		0,18		0,15	

Para toda España los resultados muestran como el porcentaje de inmigrantes en la escuela no tiene efecto estadísticamente significativo sobre el rendimiento académico en matemáticas, lectura y ciencias de los alumnos no repetidores a la edad de 15 años.

- ◆ El tamaño de la clase tiene una influencia creciente en el resultado académico a una tasa débilmente decreciente. Los óptimos se alcanzan en 24, 31 y 24 alumnos para matemáticas, lectura y ciencias respectivamente siendo tanto más negativo cuanto menos alumnos haya por clase. Al igual que en el estudio de la repetición de curso la relación entre el tamaño del grupo y el rendimiento académico no es concluyente y por tanto no se puede afirmar que menos alumnos por clase supongan *per se* un mayor rendimiento académico de los alumnos.
- ◆ El nivel socioeconómico medio de los alumnos que comparten escuela es una variable significativa en las tres pruebas de rendimiento. El paso de un alumno de un aula en el primer cuartil a un aula en el tercer cuartil le supone una ganancia de más de 24 puntos en la prueba de matemáticas, lo que supone un incremento en el rendimiento medio de aproximadamente el 4,7%.
- ◆ Los chicos obtienen resultados significativamente mejores en matemáticas y ciencias mientras que las chicas son mejores en comprensión lectora. Este resultado es consistente con la literatura en economía de la educación.
- ◆ El hecho de ser inmigrante afecta negativamente al rendimiento del propio alumno en todas las pruebas. En efecto, a pesar de no ir ningún curso retrasado los inmigrantes de 15 años obtienen un rendimiento en matemáticas un 6,3% más bajo que el de los alumnos españoles.
- ◆ Una variable con un efecto muy significativo sobre los rendimientos es la relacionada con el clima escolar percibido por el alumno. Los alumnos que declaran que en su grupo existe un mayor grado de disciplina por parte del profesor tienen menor probabilidad de ser repetidores además de tener un mejor rendimiento académico.

## Conclusiones

- ◆ El objetivo de este trabajo ha sido analizar si la presencia de inmigrantes influye en la probabilidad de que un alumno repita curso y estudiar si los alumnos españoles que no han repetido curso ven disminuir su rendimiento por la presencia de alumnos inmigrantes que tampoco han repetido curso. Para realizar este análisis se han utilizado los datos españoles del Proyecto PISA 2003 de la OCDE. Las principales conclusiones de los análisis realizados son, en primer lugar, que la integración del alumnado inmigrante ha sido resuelta por el sistema educativo mediante la política de repetición de curso con el objetivo de adecuar su nivel de partida al de conocimientos. Además, se ha constatado que la presencia de alumnos inmigrantes no afecta al rendimiento académico del resto de compañeros no repetidores en ninguna de las pruebas analizadas. Finalmente, y en concordancia con la literatura empírica previa, los resultados obtenidos muestran que el nivel socioeconómico del alumno influye significativamente en la probabilidad de ser repetidor, de manera que a mayor nivel socioeconómico la probabilidad disminuye. El nivel socioeconómico de los compañeros de escuela afecta de forma significativa tanto a la probabilidad de ser repetidor como al rendimiento académico de los alumnos no repetidores.

## Referencias

- 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.
- Dills, A. K. (2005). "Does cream-skimming curdle the milk? A study of peer effects", *Economics of Education Review*, 24, 19-28.
- Ganzeboom, H. B. G., De Graaf, P., Treiman, J. y De Leeuw, J. (1992). "A standard internacional socio-economic index of occupational status", *Social Science Research*, 21 (1), pp. 1-56.
- Hanushek, E. A. (1986). "The economics of schooling" *Journal of Economic Literature*, 24 (3), 1141-1171.
- Hanushek, E.A. (1997). "Assessing the effects of school resources on student performance: An update". *Educational Evaluation and Policy Analysis*. 19, 141-164.
- Hanushek, E. A. (2003). "The failure of input based schooling policies." *The Economic Journal*, 113, 64-98.
- 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.
- Hoxby, C. M. (1999). "The productivity of schools and other local public goods producers". *Journal of Public Economics*, 74, 1-30.
- Hoxby, C. M. (2000). "The effects of class size on student achievement: new evidence from population variation". *Quarterly Journal of Economics*, 115, 1239-1285.
- Krueger, A. B. (1999). "Experimental estimates of educational production functions". *Quarterly Journal of Economics*, 114, 497-532.
- Mislevy, R. J. (1991). "Randomization-based inference about latent variable from complex samples". *Psychometrika* 56, Psychometric Society, Greensboro, pp. 177-196.
- Mislevy, R. J., A. E. Beaton, B. Kaplan y K. M. Sheehan (1992). "Estimating population characteristics form sparse matrix samples of item responses", *Journal of Educational Measurement* 29, pp.133-161.
- OECD (1999). *Classifying educational programmes: Manual for ISCED-97 implementation in OECD countries*, Paris. OECD

- OECD (2001). Manual for the PISA 2000 Database. Organisation for Economic Co-operation and Development. Paris.
- OECD (2001). knowledge and skills for life. First results from the O. C. D. E. Program for International Student Assessment (PISA) 2000. Organisation for Economic Co-operation and Development. Paris.
- OECD (2002). PISA 2000 Technical Report. Organisation for Economic Co-operation and Development. Paris.
- OECD (2004). Learning for Tomorrow's World. First Results from PISA 2003. Organisation for Economic Co-operation and Development. Paris.
- OECD (2005a). PISA 2003 Data Analysis Manual. SAS users. Organisation for Economic Co-operation and Development. Paris.
- OECD (2005b). PISA 2003 Technical Report. Organisation for Economic Co-operation and Development. Paris.
- OECD (2005c). PISA 2003 Data Analysis Manual. SPSS users. Organisation for Economic Co-operation and Development. Paris.
- Rivkin, S. G., Hanushek, E. A. y Kain, J. F. (2005). "Teachers, schools and academic achievement". *Econometrica*, 73 (2), 417-458.