

DETERMINANTES DE LA DESIGUALDAD EDUCATIVA EN ESPAÑA

Eva Crespo Cebada

(ecreceb@unex.es)

Carlos Díaz Caro

(carlosdc@unex.es)

Jesús Pérez Mayo

(jperez@unex.es)

Resumen

La última publicación del Informe PISA 2009 ha puesto de manifiesto la existencia de diferencias notables en las puntuaciones obtenidas en las diversas pruebas de aptitudes y conocimientos por los alumnos de las regiones españolas que han participado en dicho estudio. El presente trabajo tiene por objeto abundar en la capacidad en términos de igualdad de oportunidades que ofrece el sistema educativo español. Con este propósito, se plantea un análisis de equidad a nivel de alumno donde se incluye información relativa a diferentes factores asociados con el entorno socioeconómico y familiar de los estudiantes, así como las características de los centros educativos que pueden condicionar sus resultados. Para la realización de dicho análisis se implementa la metodología propuesta por *Wagstaff et al.* (2003), que nos permite cuantificar el grado de (in)equidad horizontal en la educación secundaria en España mediante el cálculo de índices de concentración, así como identificar la contribución de distintas variables a la desigualdad en los resultados. La construcción de dichos índices nos permite dar respuesta a la capacidad distributiva del sistema de educación pública en España a partir de la contribución de los factores de no necesidad al índice de (in)equidad horizontal. En concreto, si estas variables no presentan una influencia significativa en la desigualdad de los resultados educativos, ésta podría deberse a otros factores como la motivación del alumno, el esfuerzo dedicado en el estudio, la calidad de los métodos docentes, entre otros. Los resultados obtenidos muestran que la desigualdad en la competencia académica comprensión lectora está condicionada fundamentalmente por las variables relacionadas con el entorno personal y familiar del alumno, en concreto por la desigualdad motivada por la condición de repetidor, la titularidad de la escuela, así como la formación y/o cualificación de los padres.

Palabras clave: Educación, Desigualdad, Igualdad de Oportunidades, PISA.

Clasificación JEL: I21, D63.

1. Introducción

La última publicación del Informe PISA 2009 ha puesto de manifiesto la existencia de diferencias notables en las puntuaciones obtenidas en las diversas pruebas de aptitudes y conocimientos por los alumnos de las regiones españolas que han participado en dicho estudio. La equidad en educación es un elemento esencial para la cohesión social y el crecimiento económico (Barro, 2000; Barro and Lee, 2001). Existe un gran consenso en que las diferencias en el resultado educativo atribuido al esfuerzo o habilidades innatas deben ser consideradas como inevitables y legítimas, sin embargo aquellas que incrementan la falta de oportunidad o exclusión social representan una importante pérdida de bienestar social para la sociedad.

La desigualdad en las competencias educativas atribuibles al entorno socioeconómico del estudiante está considerada como uno de los factores de estudio con mayor relevancia dentro de los países desarrollados. Existen varios motivos que ayudan a explicar la existencia de esta fuerte relación entre el entorno socioeconómico y el rendimiento académico. Entre los argumentos más extendidos destaca el peso significativo del entorno familiar del alumno, así los padres que pertenecen a un alto entorno socioeconómico podrían destinar una mayor cantidad de recursos para la formación académica de sus hijos, tanto cuantitativa como cualitativamente (Becker, 1964).

Por tanto, la composición socioeconómica de las escuelas podría ser un factor que afectara al mayor/menor grado de igualdad de oportunidades del sistema educativo. En concreto, existe evidencia empírica acerca de que las escuelas mejor posicionadas en términos de resultados son aquellas con estudiantes que pertenecen a entornos socioeconómicos más altos (Coleman y otros 1966; OECD, 2001; OECD, 2004; OECD, 2005). Entre los estudios que abordan el análisis de la desigualdad en los resultados académicos para diferentes países que participan en la base de datos PISA merecen ser destacados Ammermueller (2007), Doyle (2008), Fertiga y Smchmidt (2002) y Nash (2003). Sin embargo, la aplicación de las metodologías propuestas por Wagstaff y otros (1999) y Kakwani, Wagstaff y Van Doorslaer (1997), ampliamente utilizadas en el ámbito de la economía de la salud, han sido escasamente explotadas en el sector educativo para datos a nivel de alumno (Martins y Veiga (2010) y Oppedisano y Turati (2009).

La propia Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura (UNESCO) plantea que la equidad en educación implica educar de acuerdo a las diferencias y necesidades individuales. Esto es que las condiciones económicas, demográficas, geográficas, étnicas o de género no supongan un impedimento para el aprendizaje. Asimismo, la Ley Orgánica de Educación española, en su Título II-Equidad en la educación, versa sobre el apoyo al alumnado con necesidades especiales y la compensación de las desigualdades educativas derivadas de factores sociales, económicos, culturales, geográficos, étnicos o de otra

índole mediante políticas de educación compensatorias que refuercen la acción del sistema educativo. De esta forma, se eliminarían los problemas asociados a la desigualdad de oportunidades, no existiendo una correlación directa y positiva entre resultados y capacidad económica en el hogar. De modo que, el rendimiento académico estaría determinado por el nivel de esfuerzo y dedicación de los alumnos, así como la propia capacidad innata del alumno y las estrategias pedagógicas, pero sin influir el entorno socioeconómico de procedencia.

Por tanto los objetivos de este trabajo son múltiples. En primer lugar, se estimará un modelo multinivel que nos permita analizar cuáles son los factores que inciden en el rendimiento académico, para lo cual se contará con información tanto a nivel individual como escolar. En segundo lugar, se descompone el índice de concentración del resultado académico en relación al entorno socioeconómico, usando para ello las elasticidades obtenidas en la regresión multinivel a nivel del alumno. A partir de esta estimación es posible obtener una aproximación de los factores determinantes del rendimiento educativo, así como de la desigualdad en términos de resultados para alumnos que pertenecen a diferentes entornos socioeconómicos. A continuación identificaremos un conjunto de variables de necesidad, esto es factores que afectan al rendimiento académico per se como el sexo, la edad, la condición de repetidor y/o inmigrante, es decir aquellas características personales del alumno que pueden incidir sobre el rendimiento académico. Por otra parte, las variables relacionadas con el entorno socioeconómico de las familias y ciertas características de los centros educativos serán consideradas los factores denominados de no-necesidad, es decir que afectan a la desigualdad académica. Para ello se calcula un índice de inequidad en los resultados basándonos en la metodología propuesta por Wagstaff y otros (2003).

El trabajo está organizado de la siguiente forma. En la sección segunda se exponen las principales características de la función de producción utilizada. En la siguiente sección se desarrollan las dos metodologías que van a ser combinadas: el análisis multinivel y la descomposición del índice de desigualdad. En el apartado cuarto se describe la base de datos, PISA 2009, así como las variables utilizadas en dicho análisis. A continuación, en la sección quinta se exponen los resultados obtenidos para ambas metodologías comentadas- la estimación multinivel y la descomposición de la desigualdad- y, asimismo, se obtiene el índice de inequidad educativo. Finalmente, el trabajo concluye con un resumen de las principales conclusiones obtenidas en nuestro estudio.

2. Metodología

2.1. Función de producción educativa

La función de producción educativa presenta ciertas similitudes con una función de producción tipo, en la que los alumnos obtienen unos resultados académicos a partir de las

características personales y familiares de cada uno de ellos y de otros factores disponibles en las escuelas como las instalaciones, el equipamiento de los centros, el conocimiento de los profesores, etcétera. En este trabajo se propone la utilización de la especificación propuesta por Levin (1974) y Hanushek (1972, 1979):

$$Y_{is} = f(B_{is}, S_{is}, P_{is}, I_{is}) \quad (1)$$

siendo Y el resultado de un estudiante cualquiera i en una escuela s , cuyo valor suele estar representado por los resultados obtenidos por los alumnos en pruebas estandarizadas. Este vector de resultados depende a su vez de un conjunto de factores, entre los que destacan, el entorno socioeconómico (B_{is}), como principal característica familiar, los inputs de la escuela (S_{is}), tales como el material educativo, los profesores, las infraestructuras de la escuela, la influencia del resto de compañeros en la clases, (P_{is}), así como las propias cualidades innatas del alumno (I_{is}).

2.2. Regresión multinivel

Para la estimación de la función de producción educativa se ha seguido el enfoque de los modelos multinivel por varias razones. En primer lugar, estos modelos permiten conocer tanto el valor promedio de los factores explicativos sobre la variable dependiente como la variación que se produce entre los diferentes niveles, en este caso las escuelas. De esta forma, el modelo no estima una sola regresión, sino que calcula múltiples rectas correspondientes a cada escuela.

En segundo lugar, la estimación mediante mínimos cuadrados ordinarios (MCO) no produce errores estándar correctos, debido a que los alumnos que pertenecen a una misma escuela presentan valores iguales para las variables relativas a la escuela. De manera que, la correlación promedio entre variables de los alumnos que pertenecen a un mismo centro será superior a la que existe entre alumnos de diferentes escuelas (Hox, 1995). Así, los modelos multinivel son ampliamente utilizados en aplicaciones empíricas para el ámbito educativo, ya que los alumnos se encuentran agrupados por centros (Bryk y Raudenbush, 1992; Snijders, 1999; Bickel, 2007).

A continuación se muestra la estimación de un modelo multinivel en su forma más básica, esto es considerando tan solo dos niveles, alumnos y escuelas:

$$\begin{aligned} Y_{ij} &= \beta_{0j} + \varepsilon_{ij} \\ \beta_{0j} &= \gamma_{00} + U_{0j} \end{aligned} \quad (2)$$

donde Y_{ij} representa el resultado académico para el alumno i en la escuela j , el cual se explica según la desviación de la escuela j (u_j) respecto de los resultados medios de todas las escuelas (γ_0) y la desviación del estudiante i en relación a la media de los resultados obtenidos por los alumnos que pertenecen a su misma escuela j .

2.3. Medición de la equidad horizontal en términos de resultados académicos y descomposición de sus factores determinantes

El análisis de la equidad constituye un aspecto de gran relevancia en el ámbito de la Economía Pública, que se corresponde con su notable desarrollo en la literatura. De las dos dimensiones posibles de la equidad, vertical y horizontal, se aborda la medición de esta última. Esto es, se tratará de identificar hasta qué punto los individuos considerados “iguales” en necesidad obtienen resultados similares en la competencia lectora.

En la práctica existen condicionantes o factores socioeconómicos, como el ingreso familiar del alumno o el nivel de estudios de sus padres, la titularidad del centro en el que estudia el alumno, entre otros, que afectan al rendimiento académico. Por tanto, existirá inequidad en el sistema educativo cuando los resultados de los alumnos se encuentren determinados por factores de no necesidad, es decir, cuando el rendimiento de los mismos no se corresponde con sus propios factores de necesidad.

La aportación de *Wagstaff y otros* (2003) permite identificar y cuantificar en qué medida los distintos determinantes de la utilización de los servicios sanitarios inciden sobre la (in)equidad. Los autores proponen una nueva metodología que consiste en aplicar un índice de concentración para analizar las variaciones en la utilización de los servicios sanitarios a partir de la estimación de un modelo econométrico. Se trata de descomponer la desigualdad asociada a factores socioeconómicos en función de la contribución parcial de cada uno de los regresores de la estimación sobre el índice de concentración total, siendo dicha contribución el producto de la elasticidad de la variable dependiente en relación a cada factor explicativo y su correspondiente índice de concentración.

En nuestro caso, se utiliza el índice de concentración (IC) para medir el grado de desigualdad de los alumnos en la competencia lectora, los cuales son ordenados en función del entorno socioeconómico al que pertenece cada uno de ellos. Para ilustrarlo, supongamos una medida del resultado académico y_i , y su curva de concentración. Dicha curva relaciona la proporción acumulada de los alumnos, siendo ésta ranqueada por una variable que represente las características socioeconómicas del alumno con la proporción acumulada de los resultados académicos en lectura. De tal forma que el IC se calcula según la siguiente expresión:

$$IC = \left(\frac{2}{\bar{y}}\right) \text{cov}(\bar{y}_i, R_i) \quad (3)$$

donde $\bar{y} = E(y_i)$ representa los resultados académicos y R_i la proporción acumulada de población ordenada por la variable socioeconómica. También, será necesario normalizar el índice por la media de las variables o por los límites del IC en caso que las variables sean dicotómicas (ver *Wagstaff, 2005*).

2.4. Descomposición del índice de concentración a partir de la regresión Multinivel

En el presente trabajo se incorpora la metodología de *Wagstaff y otros (2003)* para aproximar el nivel de equidad horizontal en la competencia lectora para el caso español. Para ello necesitamos determinar qué parte de los resultados académicos puede explicarse por factores de necesidad y qué parte se debe a factores de no necesidad. En primer lugar se obtiene el modelo de producción educativa en función de un vector de características personales del alumno (factores de necesidad), así como otras relativas a la familia y a los centros escolares (factores de no necesidad), que pueden afectar a los resultados de los alumnos.

$$Y_i = \alpha + \beta x_i + \delta z_i + u_i \quad (4)$$

donde Y_i es el resultado académico en lectura de un alumno tipo i , x_i y z_i son los vectores de variables de necesidad y no necesidad, respectivamente, β y δ los coeficientes asociados a cada uno de estos vectores y u_i el término error de la regresión. Aplicando el teorema de *Rao* para la desigualdad de la renta a la expresión anterior, el índice de concentración puede expresarse como la suma ponderada de los IC de los distintos regresores, donde la ponderación corresponde a la elasticidad del rendimiento académico con respecto a cada una de las variables explicativas, como se aprecia en la Expresión 5:

$$IC = \sum_i \left(\beta_i^m \frac{\bar{x}_i}{\mu} \right) IC_i + \sum_j \left(\gamma_j^m \frac{\bar{z}_j}{\mu} \right) IC_j + \frac{GC_u}{\mu} \quad (5)$$

donde μ es la media de la variable del resultado académico, \bar{x}_i (\bar{z}_j) es la media de cada factor de necesidad (no necesidad), IC_i (IC_j) es el índice de concentración parcial para las variables

de necesidad (no necesidad) y GC_u es el índice de concentración generalizado para el término error. Este último refleja la inequidad en los resultados académicos no explicada por las variables x_i y z_j , como por ejemplo el grado de esfuerzo y la motivación del alumno.

A partir de la Expresión (5) podemos obtener el índice de inequidad horizontal (IH) en los resultados académicos como diferencia entre el índice de concentración total (IC) y el sumatorio de los índices de concentración parciales de las variables de necesidad (IC_n) (Wagstaff y van Doorslaer, 2003). De esta forma, la inequidad vendrá dada por el índice de concentración derivado de los factores denominados de “no necesidad” (IC_{nn}), tal y como muestra la siguiente expresión:

$$HI = IC - IC_n = IC_{nn} \quad (6)$$

donde valores positivos implican inequidad en términos de resultados académicos que favorece a los individuos con mejor situación socioeconómica, mientras que un signo negativo indica que son los individuos con menos recursos lo que más se benefician de dicha inequidad.

3. Datos y variables PISA 2009

La estimación de la función educativa requiere la identificación de los factores productivos que intervienen en el proceso educativo (inputs), así como de los conocimientos y habilidades adquiridos por el alumno en su paso por dicho proceso (outputs). En este trabajo se utiliza información relevante, tanto del entorno familiar del alumno como de la escuela, relativa al proyecto PISA (Programme for International Students Assessment) 2009. En dicho programa se evalúa el conjunto de competencias de los alumnos de 15 años en tres disciplinas: matemáticas, comprensión lectora y ciencias, si bien en cada oleada se analiza una de estas disciplinas con mayor profundidad, siendo comprensión lectora la disciplina central en la oleada del 2009.

Dada la dificultad para valorar los conocimientos y habilidades adquiridos en un determinado contexto y durante un curso académico, en este trabajo, y en línea con otros estudios, se aproxima el output educativo mediante el resultado obtenido en la prueba de conocimientos en lectura por los alumnos españoles que participan en la prueba PISA 2009.

Para la aplicación empírica se utilizan los 5 valores plausibles¹ para comprensión lectora proporcionados por PISA. De esta forma los analistas de PISA pretenden que se incorpore en el análisis las posibles circunstancias adversas que pueden afectar al alumno el día del examen,

¹ Los valores plausibles son valores aleatorios que se extraen aleatoriamente de la función de distribución de los resultados de cada alumno en cada disciplina evaluada.

como que se encuentre enfermo, nervioso, tenga mala suerte, entre otros condicionantes, que no son controlables por el alumno en el momento de realizar la prueba. Asimismo, la posibilidad de contar con 5 valores para cada disciplina permite realizar el análisis reiteradas veces, obviando posibles problemas de sesgo en los resultados.

En cuanto a las variables explicativas del rendimiento académico, se han considerado aquellas que participan directamente en el proceso educativo, diferenciando entre variables a nivel de individuo y de escuela.

Variables a nivel de individuo:

- SEXO: variable dummy que toma el valor uno en caso que el alumno sea chica y cero en caso de ser chico.
- EDAD: variable continua que representa la edad del alumno.
- PRIMARIA: variable dummy que toma el valor uno en caso que el alumno haya asistido a algún curso de pre-primaria y cero en caso contrario.
- REPETIDOR_1: variable dummy que toma el valor uno en caso de que el alumno haya repetido un curso y cero en caso contrario.
- REPETIDOR_2: variable dummy que toma el valor uno si el alumno ha repetido más de un curso y cero en caso contrario.
- INMIGRANTE_1: variable dummy que toma el valor uno si el alumno es inmigrante de primera generación, esto es que el alumno y/o los padres hayan nacido en el extranjero, y cero en caso contrario.
- INMIGRANTE_2: variable dummy que toma el valor uno si el alumno es inmigrante de segunda generación, de manera que alguno de los padres del alumno haya nacido en el extranjero, y cero en caso contrario.
- NUCLEOFAMILIAR: variable dummy que toma el valor uno si el alumno convive con sus padres, siendo cero en caso contrario.
- MADRE_EDUC: variable dummy que toma el valor uno si la madre del alumno se encuentra ocupada en una profesión que requiriere un elevado nivel de cualificación, siendo cero en caso contrario.
- PADRE_EDUC: variable dummy que toma el valor uno si el padre del alumno se encuentra ocupado en una profesión que requiriere un elevado nivel de cualificación, tomando el valor cero en caso contrario.
- ESCRITORIO: variable dummy que toma el valor uno en caso que el alumno disponga de un escritorio propio en el hogar, siendo cero en caso contrario.

- **HABITACIÓN:** variable dummy que toma el valor uno si el alumno dispone de habitación propia en el hogar y cero en caso contrario.
- **LUGAR_ESTUDIO:** variable dummy que toma el valor uno si el alumno dispone de un lugar de estudio propio en el hogar y el valor cero en caso contrario.
- **ORDENADOR:** variable dummy que toma el valor uno si el alumno dispone de un ordenador en el hogar, siendo cero en caso contrario.
- **LIBRO<20:** variable dummy que toma el valor uno si el alumno dispone de menos de 20 libros en el hogar, siendo cero en caso contrario.
- **LIBRO>200:** variable dummy que toma el valor uno si el alumno dispone de más de 200 libros en el hogar, siendo cero en caso contrario.

Variables a nivel de escuela

- **CLIMA:** variable continua que ofrece información acerca del nivel de disciplina en el aula para aquellos alumnos que comparten la misma escuela.
- **PRIVADA:** variable dummy que toma el valor uno si la escuela es privada, esto es que tanto la gestión como financiación son privadas, y toma el valor cero si la escuela es pública o concertada.
- **PCGIRLS:** variable continua que representa el porcentaje de chicas en la escuela. Esta variable se construye como la ratio del número de alumnas entre el número total de alumnos en la escuela.
- **SCHSIZE,** representa el tamaño de la escuela, expresado como el número total de alumnos dentro de un mismo centro.
- **STRATIO:** variable continua que representa la ratio alumno-profesor, como el cociente entre el número total de alumnos entre el número total de profesores en cada escuela.
- **IRAPTCOMP:** variable continua que representa el porcentaje de ordenadores por alumno, como el cociente entre el número total de ordenadores y el tamaño de la escuela.
- **SCMATEDU:** representa la calidad de los recursos escolares. Dicha variable es construida a partir de las respuestas de los directores de cada centro a un conjunto de preguntas relacionadas con la disponibilidad de recursos didácticos, como son: software educativo, calculadoras, libros, recursos audiovisuales y equipo de laboratorio.

4. RESULTADOS

En este apartado se presentan los resultados obtenidos una vez realizada la regresión multinivel, así como la descomposición del índice de concentración de los resultados educativos en los diferentes factores que explican la variable dependiente.

En concreto, en primer lugar se presenta el resultado de la regresión multinivel. La Tabla 1 recoge los resultados de tres modelos. El primer modelo (Modelo 1) estima de forma restringida, es decir incluyendo la constante como única variable. Tal y como muestra la Tabla 1, en el caso del modelo restringido la varianza de los resultados es explicada en su mayoría por la varianza atribuible al alumno (80%), mientras que el peso atribuible a la escuela es tan sólo el 20%. El segundo modelo (Modelo 2) incorpora las variables individuales además de la constante, en cuyo caso la variación de varianza atribuible a la escuela es incluso menor que en el modelo anterior, pasando a ser un 16% frente al 84% de varianza atribuible al alumno. Finalmente, el último modelo (Modelo 3) se trata de una estimación completa con todas las variables, tanto las relativas al alumno como a la escuela. Por último, en relación a la descomposición de la varianza del término error, dicha estimación arroja resultados similares al modelo 2, siendo la proporción de varianza explicada por la escuela del 15% en este caso.

Tabla I: Regresión multinivel para los modelos reducido y ampliados

	<i>Modelo 1</i>	<i>Modelo 2</i>	<i>Modelo 3</i>
<i>Varianza Individuo</i>	8.234,60	5.715,66	5.715,33
<i>Varianza Escuela</i>	2.024,01	1.079,86	1.029,58
<i>Varianza</i>	10.258,61	6.795,53	6.744,90
<i>% Individuo</i>	80,27%	84,11%	84,74%
<i>% Escuela</i>	19,73%	15,89%	15,26%

Fuente: elaboración propia a partir de datos de PISA 2009

La Tabla 2 muestra los resultados del análisis multinivel para cada uno de los modelos mencionados anteriormente. En cada modelo pueden distinguirse tres columnas, el coeficiente, la desviación estándar y el p-valor, respectivamente. Asimismo, un signo positivo en el coeficiente indica que la variable correspondiente influye positivamente en el resultado académico en lectura, mientras que un signo negativo indica justo lo contrario. A continuación se comentan los resultados obtenidos de la estimación de cada modelo, tal y como se aprecia en la Tabla 2.

La incorporación de las variables de tipo individual en el modelo 2 pone de manifiesto el marcado efecto negativo y significativo que presenta la condición de repetidor, especialmente si el alumno ha repetido en más de una ocasión, sobre los resultados en la prueba de

conocimientos en lectura. Asimismo, la condición de inmigrante, tanto de primera como de segunda generación, también afecta de forma negativa al rendimiento académico en lectura, si bien este efecto únicamente es significativo para los inmigrantes de primera generación. A su vez, las variables que recogen información personal del alumno muestran un marcado efecto positivo y significativo sobre el rendimiento en lectura, en concreto destacan la condición de chica, la edad y haber realizado un curso en la etapa anterior a la enseñanza primaria. Por otra parte, los factores asociados al entorno familiar suponen en su mayoría un efecto positivo sobre el rendimiento en lectura, a excepción de poseer una cantidad de libros en el hogar inferior a 25, cuya influencia es negativa y significativa, y que el alumno posea habitación propia, ya que el resultado no es significativo.

Sin embargo, la posesión de ordenador propio en el hogar, así como que los padres tengan un trabajo que requiera un elevado nivel de cualificación, son factores que influyen positiva y significativamente en el rendimiento de la disciplina lectura. Por último, la incorporación de variables de ámbito individual como a nivel de escuela, modelo 3, arroja unos resultados muy similares a los del modelo 2 en cuanto a las variables individuales se refiere. En lo que respecta a las variables a nivel de escuela, los resultados parecen indicar que los alumnos que pertenecen a centros de gestión privada obtienen significativamente mejores resultados en lectura. La ratio alumno-profesor también parece ser un factor que favorece al rendimiento académico. No obstante, el tamaño del centro escolar, así como la calidad de los recursos escolares, el clima de disciplina en el aula o la ratio de ordenadores por alumno en el aula no son factores que afecten significativamente al rendimiento en lectura. Estos resultados coinciden con los alcanzados en otros estudios como Calero y Escardibul, (2007), Calero y otros (2009), Mancebón y otros (2010), Salinas y Santín (2012), Cordero y otros (2010), Cordero y otros (2011), Perelman y Santín (2011).

TABLA II: Modelos de regresión multinivel

Variables	<i>Modelo 1</i>			<i>Modelo 2</i>			<i>Modelo 3</i>		
	<i>Coef.</i>	<i>Est.Dev</i>	<i>P-Valor</i>	<i>Coef.</i>	<i>Est.Dev</i>	<i>P-Valor</i>	<i>Coef.</i>	<i>Est.Dev</i>	<i>P-Valor</i>
Constante	485,240	1,620	0,000	310,120	27,090	0,000	320,130	27,110	0,000
<i>NIVEL ALUMNO</i>									
Sexo				22,853	0,960	0,000	22,670	0,970	0,000
Edad				7,739	1,680	0,000	7,650	1,680	0,000
Primaria				16,680	2,110	0,000	17,140	2,120	0,000
Repetidor_1				-66,808	1,240	0,000	-66,400	1,250	0,000
Repetidor_2				-121,127	1,860	0,000	-120,350	1,870	0,000
INMIG_1				-10,798	2,000	0,000	-10,980	2,010	0,000
INMIG_2				-4,520	4,450	0,222	-5,430	4,450	0,222
Nuleo_Familiar				3,506	1,340	0,062	2,580	1,350	0,056
Madre_Cualificada				10,383	1,200	0,000	10,050	1,210	0,000
Padre_Cualificado				8,408	1,160	0,000	7,950	1,170	0,000
Escritorio				16,506	3,390	0,000	9,010	3,400	0,008
Habitación				1,056	1,480	0,863	0,270	1,490	0,855
Lugar_estudio				11,766	1,920	0,000	10,640	1,920	0,000
Ordenador				20,331	2,050	0,000	18,080	2,060	0,000
Libros < 25				-27,253	1,350	0,000	-27,650	1,350	0,000
Libros >200				10,792	1,170	0,000	11,370	1,170	0,000
<i>NIVEL ESCUELA</i>									
Clima_Aula							0,000	0,000	0,000
Titularidad							9,530	2,530	0,000
Pcgrils							-0,050	0,320	0,114
Schsize							0,010	0,000	0,111
Stratio							0,000	0,000	0,005
Iratcomp							0,000	0,000	0,890
Scmatedu							0,000	0,000	0,861

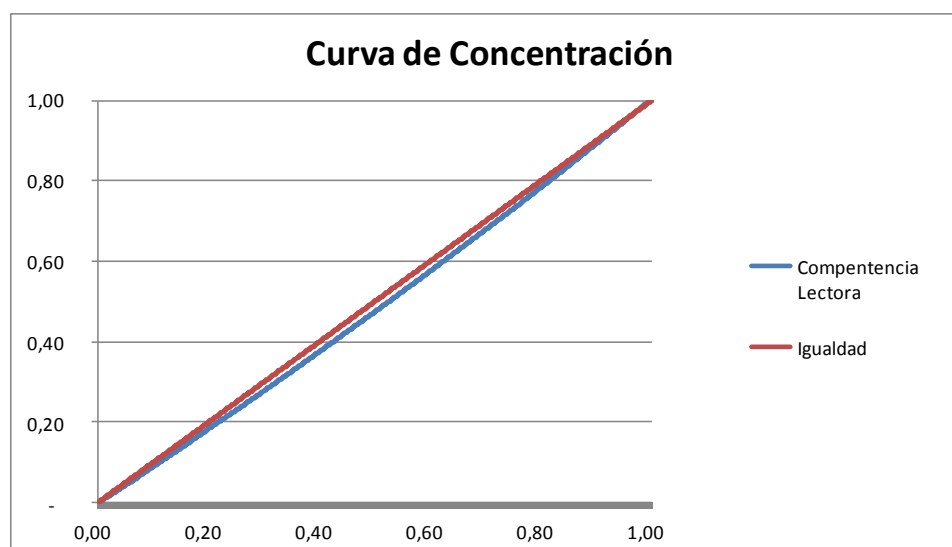
Fuente: elaboración propia a partir de datos de PISA 2009

4.1. Descomposición del índice de concentración.

Una vez se ha analizado cómo afectan las distintas variables sobre el rendimiento académico, podemos obtener una medida de la equidad horizontal mediante el cálculo del índice de concentración descrito en la sección anterior.

En la gráfica que aparece a continuación se representa la desigualdad en términos de resultados. El eje horizontal representa la proporción acumulada de individuos ordenados por su condición socioeconómica (*ESCS*), y en el eje vertical la proporción acumulada de los resultados de los alumnos en la competencia lectura. El IC toma valores entre $[-1, 1]$ y su valor se corresponde con el doble de la superficie comprendida entre la diagonal y la curva de concentración. Los valores negativos del IC indican que existe desigualdad a favor de los individuos más desfavorecidos económicamente, mientras que un signo positivo implica justamente lo contrario, que los individuos con mayor renta acaparan un porcentaje superior de los resultados académicos. A su vez, un signo negativo (positivo) del IC se corresponde con una curva del IC por encima (debajo) de la diagonal. En el caso que el IC tomara el valor nulo las distribuciones de los resultados y del estatus socioeconómico de los individuos serían iguales, por lo que habría igualdad perfecta en términos de resultados y la curva del IC coincidiría con la diagonal.

GRÁFICO I: Curva de Concentración



Fuente: elaboración propia a partir de datos de PISA 2009

Además, podemos obtener una medida de la inequidad horizontal en términos de resultados a partir de la descomposición del IC entre variables de necesidad y no necesidad

(Wagstaff y otros, 2003). Entre los factores de necesidad hemos considerado aquellos relativos al alumno -edad, sexo, condición de inmigrante y repetidor- mientras que el resto de variables hacen referencia al nivel socioeconómico o la situación familiar, y de las escuelas. En el Tabla 3 se presenta el IC total y los IC parciales, correspondiendo la tercera fila con el índice de inequidad horizontal.

TABLA III: Índice de inequidad horizontal en el sistema educativo español

	<i>Media</i>	<i>Desv. Estd.</i>	<i>95% Int. Conf.</i>		<i>Porcentaje del IC</i>
<i>IC</i>	0,0362	0,0011	0,0362	0,0363	100,00%
<i>ICn</i>	0,0146	0,1423	0,0129	0,0164	40,38%
<i>HI=ICnn</i>	0,0158	0,2397	0,0129	0,0187	43,66%
<i>GCu</i>					15,95%

Fuente: elaboración propia a partir de datos de PISA 2009

Los resultados muestran que el IC del sistema educativo es 0,0362, lo cual indica que existe desigualdad en términos de resultados a favor de los individuos de mayor nivel socioeconómico. Comprobamos también que dicha desigualdad se debe, fundamentalmente, a diferencias en los factores de no necesidad; en concreto, el 43,66% de la desigualdad se atribuye a diferencias en los factores del entorno familiar y de la escuela, mientras que el 40,38% se debe a variables del ámbito personal del alumno. Asimismo, apelando al Teorema Central del límite para estimar la significatividad, tanto del índice de concentración total como de los índices parciales, podemos descartar la hipótesis de que estos índices sean significativamente distintos de cero, es decir que los resultados mostrados en la Tabla 3 son significativos al 95%.

A continuación se muestran los resultados de la descomposición de la desigualdad en los resultados académicos siguiendo la metodología expuesta anteriormente. En primer lugar, merece destacarse que, el índice de concentración parcial de cada variable representa cómo afecta la distribución de dicha variable a los resultados académicos en la competencia lectora, mostrando una clara desigualdad para aquellos alumnos con mayores recursos económicos en el hogar, lo cual indicaría desigualdad pro-ricos en el logro académico.

Tabla IV: Descomposición de la desigualdad de los resultados en lectura

<i>Variables</i>	<i>Efecto</i>				<i>%</i>	
	<i>Media</i>	<i>Parcial</i>	<i>Elasticidad</i>	<i>IC</i>	<i>Contribución</i>	<i>Contribución</i>
<i>Sexo</i>	0,4924	22,6690	0,0231	-0,0073	-0,0002	-0,46%
<i>Edad</i>	15,8613	7,6542	0,2507	0,0002	0,0001	0,15%
<i>Primaria</i>	0,9408	17,1402	0,0333	0,0101	0,0003	0,93%
<i>Repetidor_1</i>	0,2293	-66,3987	-0,0314	-0,1934	0,0061	16,79%
<i>Repetidor_2</i>	0,0878	-120,3474	-0,0218	-0,3632	0,0079	21,87%
<i>INMIG_1</i>	0,0741	-10,9806	-0,0017	-0,2322	0,0004	1,08%
<i>INMIG_2</i>	0,0120	-5,4330	-0,0001	-0,0953	0,0000	0,04%
<i>Nuleo_Familiar</i>	0,8425	2,5763	0,0045	0,0172	0,0001	0,21%
<i>Madre_Cualificada</i>	0,2633	10,0468	0,0055	0,5319	0,0029	8,02%
<i>Padre_Cualificado</i>	0,3144	7,9525	0,0052	0,5070	0,0026	7,22%
<i>Escritorio</i>	0,9741	9,0121	0,0181	0,0058	0,0001	0,29%
<i>Habitación</i>	0,8680	0,2716	0,0005	0,0191	0,0000	0,03%
<i>Lugar_estudio</i>	0,9216	10,6353	0,0202	0,0169	0,0003	0,95%
<i>Ordenador</i>	0,9312	18,0784	0,0348	0,0262	0,0009	2,52%
<i>Libros < 25</i>	0,2059	-27,6475	-0,0118	-0,3936	0,0046	12,77%
<i>Libros >200</i>	0,2860	11,3710	0,0067	0,3992	0,0027	7,40%
<i>Clima_Aula</i>	124,8206	-0,0046	-0,0012	0,3507	-0,0004	-1,14%
<i>Privada</i>	0,3876	9,5325	0,0076	0,2294	0,0018	4,83%
<i>Pcgrils</i>	48,5163	-0,0502	-0,0050	0,0568	-0,0003	-0,79%
<i>Schsize</i>	504,0908	0,0051	0,0053	0,0657	0,0003	0,96%
<i>Stratio</i>	754,0287	-0,0020	-0,0030	-0,0444	0,0001	0,37%
<i>Iratcomp</i>	479,8713	-0,0001	-0,0001	-0,0543	0,0000	0,01%
<i>Scmatedu</i>	245,6366	-0,0002	-0,0001	-0,0239	0,0000	0,01%
<i>Residuo</i>					0,0058	15,95%
<i>Total</i>					0,0362	100,00%

Fuente: elaboración propia a partir de datos de PISA 2009

La Tabla 4 muestra en la primera y segunda columnas el valor medio y el efecto parcial de cada variable. La tercera columna representa la elasticidad de los resultados con respecto a cada variable explicativa. A la vista de los resultados se aprecia que los individuos con edades más altas (*Edad*) presentan una mayor propensión al rendimiento académico. En lo que respecta a los factores socioeconómicos, se comprueba que las características familiares del alumno, como tener escritorio propio (*Escritorio*) y ordenador (*Ordenador*), lugar de estudio (*Lugar_estudio*) en el hogar o la cualificación de los padres, influyen positivamente en los

resultados en la competencia lectora, a excepción de la posesión de menos de 25 libros en el hogar (Libros <25). Respecto a los factores escolares, se aprecia que, únicamente la titularidad privada del centro y el tamaño del colegio afectan positivamente al rendimiento de dicha disciplina. La cuarta columna de dicho cuadro muestra el índice de concentración parcial (*Ick*) para cada una de las variables explicativas. El signo negativo (positivo) indica que la variable en cuestión presenta una distribución pro-pobres (pro-ricos). De estos datos se deduce fácilmente que todos los indicadores relacionados con la condición de repetidor o inmigrante se concentran relativamente más en los individuos de menor capacidad económica. Por otra parte los indicadores de facilidades en el hogar, como tener un ordenador, un lugar de estudio, escritorio y habitación propios para el alumno, así como la posesión de libros en el hogar en una cantidad superior a los 200 ejemplares, presentan una distribución pro-ricos. Entre los factores escolares, se observa que la titularidad privada, la disciplina en el aula, la proporción de chicas o el tamaño del centro presentan una distribución pro-ricos, mientras que la distribución será pro-pobres para las variables: ratio alumno-profesor, la calidad de los recursos escolares y la ratio de ordenadores por alumno.

Finalmente, en la cuarta y quinta columnas se detalla, respectivamente, la contribución absoluta y porcentual de cada uno de los componentes del IC. Una contribución absoluta negativa (positiva) de una variable supone que dicho factor explica de forma negativa (positiva) la distribución de los resultados académicos pro-pobres. Como se observa en la Tabla 4, todos los factores, tanto de necesidad como de no necesidad, presentan en su mayoría signo positivo - a excepción del sexo, la disciplina en el aula y la proporción de chicas- por lo que contribuyen a explicar la inequidad en los resultados académicos en lectura a favor de los individuos de mayor renta. La variable más relevante en la explicación de la desigualdad es la condición de repetidor, especialmente en caso que el alumno haya repetido más de un curso académico, con una contribución del 16,79% y el 21,87%, respectivamente. En relación a los factores de ámbito familiar, destaca la contribución de la cualificación de los padres, contribuyendo ambas variables alrededor del 15% al IC total, así como la posesión de libros en el hogar, que supone un 20,17% del IC total. Entre los factores relativos a la escuela, merece destacarse la contribución de la titularidad privada del centro educativo, que supone un 4,87% del IC total. Asimismo, merece ser destacado que los factores que contribuyen en un mayor porcentaje a explicar el IC total son aquellos referentes al ámbito individual, esto es las propias características personales y familiares del alumno cuya participación sería en torno al 80%, siendo únicamente el porcentaje atribuible a factores escolares entorno al 4%, mientras que el 16% restante sería el error de la estimación.

5. CONCLUSIONES

En el presente trabajo se analiza los factores determinantes en la desigualdad del rendimiento educativo en España para la disciplina comprensión lectora. Para ello, en primer lugar, se han identificado los factores explicativos del rendimiento académico de los alumnos en dicha disciplina mediante un análisis multinivel. Asimismo, se ha calculado el grado de (in)equidad horizontal en el sistema educativo español y se ha procedido a descomponer la inequidad en sus factores determinantes a partir de la metodología propuesta por *Wagstaff y otros* (2003), de modo que pudiera conocerse cuál es la contribución relativa de cada uno de ellos a la desigualdad total.

Del análisis multinivel se concluye que, los principales factores determinantes del rendimiento educativo para la disciplina comprensión lectora son, los correspondientes al ámbito personal y familiar del alumno. Entre ellos, son especialmente relevantes el sexo, así como la condición de inmigrante y repetidor, especialmente para los inmigrantes de primera generación y los alumnos que repiten más de un curso académico. Por otra parte, destaca la escasa aportación de las variables representativas de los centros. El análisis de desigualdad muestra que existe cierta inequidad en el sistema educativo español que favorece a los alumnos que proceden de un entorno socioeconómico más alto. Dicho análisis indica que el índice de concentración de los resultados en lectura alcanza un valor de 0,0362. Los factores de necesidad contribuyen a explicar dicha desigualdad en un 43,66%, frente al 40,38% que corresponde al conjunto de factores relacionados con el ámbito personal y familiar del alumno. Las variables más relevantes para explicar la desigualdad son, la condición de repetidor, con una contribución global del 38,66%, y la cualificación de los padres, con una contribución global del 15,24%. Las variables relativas a la posesión de libros en el hogar contribuyen alrededor del 20,17% a la desigualdad total. Por último, las variables representativas de facilidades en el hogar, como la posesión de escritorio propio, habitación o lugar de estudio muestran una escasa contribución al conjunto de la desigualdad (0,29%, 0,03% y 0,95%, respectivamente). Asimismo, los factores escolares parecen tener una influencia menor en la explicación de la desigualdad, siendo sólo los centros privados los que presentan una contribución mayor al IC total (4,83). Este hecho parece relevante desde la perspectiva de la política educativa, ya que los recursos destinados a la enseñanza, como el tamaño de los centros o del aula no parecen influir significativamente en la desigualdad de resultados entre alumnos de diferentes entornos socioeconómicos. Por otra parte, el destacado peso que supone la condición de repetidor en la explicación de la desigualdad de los resultados supone una llamada de atención a la política de repetición seguida en España. En este sentido, parece más adecuado la incorporación de medidas de atención temprana, así como clases de apoyo que eviten al alumno llegar a la

situación de repetición de curso. Este resultado coincide con los obtenidos en otros estudios Cordero y otros (2010), Cordero y otros (2011).

Por su parte, el índice de inequidad horizontal (0,0158) también apunta la presencia de una inequidad “pro-ricos”, estadísticamente significativa. La concentración de resultados superiores sobre los alumnos procedentes de entornos de rentas más altas en España se explica tanto por variables de necesidad (40,38%) como de no necesidad (43,66%). Sin embargo, la reducción de la desigualdad en el sistema educativo español sigue siendo un objetivo a alcanzar, en la medida en que los factores socioeconómicos y otros determinantes escolares contribuyen en un 59,62% a explicar las diferencias en resultados.

Referencias bibliográficas:

- Ammermueller A. (2007): “Poor Background or Low Returns? Why Immigrant Students in Germany Perform so Poorly in the Programme for International Student Assessment” *Education Economics*, Vol. 15, No. 2, pp. 215-230.
- Bickel, R. (2007): *Multilevel Analysis for Applied Research, It’s Just Regression*, Guilford Press, New York.
- Barro, R (2000): Inequality and growth in a panel of countries. *Journal of Economic Growth*, 5 (1), pp 5-32
- Barro, R. y Lee, J. (2001): International data on educational attainment updates and implications. *Oxford Economic Papers*, 53, pp 541-563
- Becker, G (1964): *Human capital*. New York: Columbia University Press for the National Bureau of Economic Research.
- Bryk, A.S. y Raudenbush, S.W. (1988): “Toward a More Appropriate Conceptualization of Research on School Effects: A Three-Level Hierarchical Linear model”, *American Journal of Education*, 97 (1), pp 65-108.
- Coleman, J. , Campbell, E., Hobson, C., McPartland, J., Mood, A., Weinfeld, F., et al (1966): *Inequality of educational opportunity*. Washington, DC: Office of Education, U.S. Department of Health, Education, and Welfare, U.S. Government Printing Office
- Calero J., Choi, A. y Waisgrais, S. (2009): “ Determinantes del rendimiento educativo del alumnado de origen nacional e inmigrante”, *Cuadernos Económicos del ICE*, 78:281-311
- Cordero, J.M., Crespo, E. y Santín, D. (2010): “Factors Affecting Educational Attainment: Evidence from Spanish PISA 2006 Results”, *Regional and Sectoral Economic studies*, vol. 10- 2, pp 55-76.

- Cordero, J.M., Crespo, E., Pedraja, F. y Santín, D. (2011): “Exploring Educational Efficiency Divergences Across Spanish Regions in PISA 2006”, *Revista de Economía Aplicada*, vol 57., pp 117-145
- OECD, (2001): *Knowledge and skills for life: First results from the OECD programme for international student assessment (PISA) 2000*. Paris: OECD Publishing.
- OECD, (2004): *Learning for tomorrow’s world first – results from PISA 2003*. Paris: OECD Publishing.
- OECD, (2005): *PISA 2003, technical report*. Paris: OECD Publishing.
- Fertig, M., y Smchmidt, C. (2002): *The role of background factors for reading literacy: Straight national scores in the PISA*. Mimeo, Universitat Heidelberg, IZA Bonn and CEPR London
- Hox, J. (1995): *Applied Multilevel Analysis*. Amsterdam, TT-Publikaties.
- Nash, R., (2003): Is the school composition effect real? A discussion with evidence from the UK PISA data. *School Effectiveness and School improvement*, 14(4), pp 441-457
- Kakwani, N., Wagstaff, A., y van Doorslaer, E. (1997): Socioeconomic inequalities in health: Measurement, computation and statistical inference, *Journal of Econometrics*, 77(1), pp 87-104
- Martins, L., y Veiga, P. (2010): Do inequalities in parents’ education play an important role in PISA students’ mathematics achievements test score disparities?, *Economic of Education Review*, 29, pp 1016-1033
- Oppedisano y Turati (2009). What are the causes of educational inequalities and of their evolution over time in Europe? Evidence from PISA , *Documents de Treball de l’IEB* 2011/1
- Perelman, S. y Santín, D. (2011): "Imposing monotonicity on outputs in parametric distance function estimations," *Applied Economics*, Taylor and Francis Journals, vol. 43(30), pages 4651-4661.
- Salinas, J. y Santín, D (2012): “Selección escolar y efectos de la inmigración sobre los resultados académicos españoles en PISA 2006”, *Revista de Educación*, en prensa
- Snijders, T.A.B. (1999): *Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling*, Sage Publications, London.
- Wagstaff, A. y van Doorslaer, E. (2000): “Measuring and testing for inequity in the delivery of health care”, *The Journal of Human Resources*, 35, 4: pp 716–733.
- Wagstaff, A.; van Doorslaer, E. y Watanabe, N. (2003): “On Decomposing the Causes of Health Sector inequalities, with an Application to Malnutrition Inequalities in Vietnam”, *Journal of Econometrics* 112, 1: pp 219–27.