



DOCUMENTOS DE TRABAJO

Del PROYECTO PICT 2003 N° 14534

FONCYT / SECYT

DOCUMENTO DE TRABAJO N° 4 / 2006

**Diferencias en el efecto de los factores
socioeconómicos del hogar y las escuelas en
el rendimiento escolar entre regiones
argentinas**

*Héctor Gertel, Roberto Giuliadori, Verónica
Herrero, Diego Fresoli, María Luz Vera y
Guadalupe Morra*

Octubre 2006

INSTITUTO DE ECONOMÍA Y FINANZAS

FACULTAD DE CIENCIAS ECONÓMICAS

UNIVERSIDAD NACIONAL DE CÓRDOBA

**Diferencias en el efecto de los factores socioeconómicos del hogar
y las escuelas en el rendimiento escolar entre regiones argentinas**

**Héctor Gertel, Roberto Giuliadori, Verónica Herrero,
Diego Fresoli, María Luz Vera y Guadalupe Morra**

**Instituto de Economía y Finanzas
Facultad de Ciencias Económicas
Universidad Nacional de Córdoba**

Septiembre de 2006

Para ser presentado en el XVII Encuentro Estado de la Investigación Educativa “Educación y Pobreza: Alumnos, docentes e instituciones”, Octubre de 2006.

Diferencias en el efecto de los factores socioeconómicos del hogar y las escuelas en el rendimiento escolar entre regiones argentinas

Héctor Gertel, Roberto Giuliadori, Verónica Herrero,
Diego Fresoli, María Luz Vera y Guadalupe Morra

Facultad de Ciencias Económicas
Universidad Nacional de Córdoba
[hgertel@eco.unc.edu.ar](mailto:hgartel@eco.unc.edu.ar)

I. Introducción

La presencia de una alta variabilidad en los resultados obtenidos por los estudiantes que participan de las pruebas estandarizadas anuales del Ministerio de Educación de la Nación plantea, en la Argentina, algunos interrogantes iniciales acerca de las razones de la misma. Identificar y cuantificar los factores determinantes de esta variabilidad constituye una preocupación permanente para múltiples disciplinas vinculadas con la investigación educativa, al tiempo que suponen un conocimiento clave para el diseño, la ejecución y el monitoreo de diferentes políticas educativas y presupuestarias. El objetivo de este trabajo consiste en explorar hasta que punto esta variabilidad en los resultados individuales pudiera ser atribuible a la presencia de diferencias en el nivel (o fuerza) con que se expresa el efecto de los factores socioeconómicos del hogar, por un lado, y los asociados con el trabajo en el aula y las escuelas, por el otro, en las diferentes regiones del país.

En un trabajo reciente, Morduchowicz observó que “es muy probable que aquellos alumnos pertenecientes a estratos sociales más bajo presenten un desempeño más pobre con relación a los estudiantes de estratos más altos” (Morduchowicz, 2003, pp. 45). En general, la mayoría de los trabajos que han estudiado la relación entre pobreza y resultados del aprendizaje coincidiría con esta afirmación. Sin embargo, la asociación entre ambos fenómenos dista de ser perfecta, en particular, en referencia a la Argentina, se observan marcadas diferencias regionales en el nivel del efecto atribuible al nivel de pobreza/riqueza relativa del hogar sobre los resultados del aprendizaje. También debe examinarse el papel de los recursos áulicos y de la organización escolar en los logros del aprendizaje. De otra manera, se estaría corriendo el peligro de llegar a invalidar “a-priori” cualquier potencial contribución que una reorganización del trabajo escolar (reforma educativa) podría aportar a la igualación de oportunidades en la sociedad. Estudiar las implicancias del nivel socioeconómico sobre el logro educativo desde una perspectiva de equidad resulta, entonces, clave.

Considerando esta problemática, el objetivo del presente trabajo es comparar el rendimiento en Lengua de los alumnos de 6º grado en las pruebas estandarizadas desarrolladas durante el Operativo Nacional de Evaluación 2000 (ONE/2000).

Para analizar este problema los estudios en Economía de la Educación han hecho uso tradicionalmente de la parábola de la función de producción. Así, a través de la analogía con un proceso productivo, es posible interpretar los logros obtenidos en las pruebas como un resultado importante del proceso escolar desarrollado en un período específico (el año lectivo, por ejemplo), y este proceso escolar, como un procedimiento pautado, que solo es posible realizarlo con la combinación del esfuerzo de los alumnos (con su aporte de experiencias y habilidades adquiridas con anterioridad, en la escuela y en el hogar), del trabajo docente en el aula (cuya efectividad está ligada con el nivel educativo del docente, su experiencia y su capacitación en servicio), con los recursos áulicos disponibles y con un marco de gestión específico. La aplicación de este marco conceptual no resulta sencillo, ya que en el mismo es necesario considerar numerosos aspectos correspondientes al propio alumno, de los docentes y de la escuela y sus principales interacciones.

Entre los aspectos que es posible aproximar estadísticamente, este trabajo considera los siguientes: (i) referidos a los alumnos, sus habilidades, su capacidad para el desarrollo de ciertas actividades intelectuales, su interés por una determinada asignatura, su carácter de alumno repitente, entre otras; como también factores correspondientes a su entorno de origen, el nivel educativo de los padres, el nivel socioeconómico del hogar, clima educativo del hogar; (ii) los aspectos directamente vinculados con el aula y la escuela también son considerados, las características edilicias, la disponibilidad de materiales didácticos destinados a la enseñanza de cada disciplina, y también características personales del docente, como su formación y entrenamiento. La parábola de la función de producción aporta el marco de análisis para, dada la evidencia obtenida de diferentes medidas del producto escolar, analizar cuáles de los factores mencionados resulta relevante y en qué medida opera su aporte al aprendizaje.

El estudio de los aportes de cada insumo en la función de producción se efectuó en base a los resultados de la estimación de un modelo lineal multivariable. Para la estimación se recurrió a la técnica de análisis jerárquico, o de multinivel (Hox, 1995), ya que la misma está diseñada para enfrentar el problema de “anidamiento de los datos” o falta de independencia estadística en las observaciones del nivel de rendimiento que surge cuando los alumnos se agrupan en aulas, las aulas en escuelas, y las escuelas en jurisdicciones. Los modelos tradicionales consideran a los estudiantes como observaciones independientes sin tratamiento explícito del efecto anidamiento.

En particular en este estudio analizamos la fuerza explicativa de cada uno de los factores, y si ésta difiere entre regiones. Surgen en este sentido interrogantes vinculados no sólo con el coeficiente estimado en cada una de las funciones de producción educativas regionales, sino también con los niveles promedio que alcanzan los diferentes insumos.

El trabajo discute, en la sección II, las características generales del relevamiento y la muestra de trabajo; el modelo estadístico a utilizar es presentado en la sección III, y en la sección IV se analizan los resultados obtenidos. Finalmente, se presentan las conclusiones más relevantes vinculadas con los objetivos planteados

Cabe destacar que quedan fuera del análisis consideraciones relacionadas con la calidad y la homogeneidad entre regiones de los insumos considerados: no hay evidencia acerca de si por ejemplo la posesión de materiales didácticos en diferentes zonas hace referencia al mismo tipo de elemento pedagógico, o si cuando nos referimos a cierta cantidad de años de escolarización de los docentes implica la misma calidad de formación. Estas diferencias, con cierto grado de concentración en determinadas regiones, también podrían determinar discrepancias importantes en los resultados entre las respectivas regiones, no detectables con este tipo de análisis.

II. Características del relevamiento y de la muestra de trabajo

II.1 Operativo Nacional de Evaluación de la Calidad Educativa

El Operativo Nacional de Evaluación comenzó a funcionar en el año 1993, en el que se evaluó a los alumnos del último año de los niveles primario y medios en Matemática y Lengua. El objetivo era lograr un sistema de información confiable que permitiera no sólo el desarrollo de políticas educativas, sino además un instrumento para mejorar la gestión institucional.

Desde entonces dicho operativo es realizado por el Ministerio de Cultura y Educación de la Nación Argentina. La generación y recolección de información se ha efectuado a partir de la realización anual de pruebas estandarizadas a través de evaluaciones en distintas disciplinas (Lengua, Matemática, Ciencias Sociales y Ciencias Naturales), y de cuestionarios complementarios de auto-respuestas suministrados a los alumnos, docentes y directores de escuelas. Estos últimos permiten recabar datos acerca del aprendizaje de los alumnos y su historia académica, la expectativas y motivación de los mismos, su entorno familiar; también aspectos referidos a los docentes y la organización de su trabajo, así como de la escuela en su conjunto.

La información obtenida se presenta en tres bases de datos independientes, una para cada nivel según el rol en el proceso productivo: alumnos, docentes y directores; éstas pueden ser combinadas a partir de variables de vinculación o identificación (sección y escuela a la que concurre el alumno; sección por la que responde el docente y establecimiento al que pertenece el director).

II.2 Descripción de la muestra de trabajo

Los datos provienen del Operativo Nacional de Evaluación correspondiente al año 2000. En el año elegido en el presente trabajo se realizó el operativo en forma de censo, esto es, abarcando toda la población estudiantil que cursaba el último año del EGB en la Argentina¹. Sin embargo, no se obtuvieron datos para la provincia de Neuquén y se excluyeron los alumnos de establecimientos rurales con secciones de menos de 5 alumnos.

En el presente trabajo se analizan los alumnos que concurren a establecimientos en áreas urbanas y las aulas con por lo menos 4 alumnos que contestaron todas las preguntas relevantes a este estudio. Los resultados se presentan para 7 regiones en que se ha dividido el país, de acuerdo a la descripción siguiente.

Regiones	Composición
Ciudad de Buenos Aires	Ciudad de Bs. As.
Partidos GBA	Partidos del Conurbano Bonaerense
Cuyo	Mendoza, San Juan y San Luis
Noreste	Corrientes, Formosa, Chaco y Misiones
Noroeste	Catamarca, Jujuy, La Rioja, Salta, Sgo. del Estero y Tucumán
Pampeana	Resto de Bs. As., La Pampa, Córdoba, Entre Ríos y Santa Fe
Patagonia	Chubut, Neuquén, Río Negro, Santa Cruz y Tierra del Fuego

¹ En los años restantes en los que se realizó el ONE sólo se cuenta con una muestra de la población educativa total.

Las regiones coinciden con las propuestas por el INDEC, excepto que la región “Gran Buenos Aires” se presenta desagregada en “Ciudad de Buenos Aires” y “Partidos del Conurbano Bonaerense”, dada la gran disparidad existente entre ambas.

La Tabla I presenta los datos de alumnos y secciones según gestión pública y privada por regiones. En la parte superior se presentan los valores correspondientes a la población, mientras en la parte inferior se resumen los relativos a la muestra.

Tabla I - Población y muestra de trabajo, desagregada en establecimientos públicos y privados

Población

Establecimiento	Ciudad de Bs. As.	Partidos GBA	Cuyo	Noreste	Noroeste	Pampeana	Patagonia	Totales
Público								
Alumnos	20,794	104,763	39,160	56,068	68,289	158,709	22,933	470,716
%	56.64	67.79	86.26	88.9	86.47	75.97	87.15	76.67
Secciones	976	3,895	1,860	3,086	3,347	7,215	1,071	21,450
%	59.88	67.65	88.28	91.33	88.73	78.96	87.64	79.43
Privado								
Alumnos	15,917	49,779	6,240	7,000	10,678	50,162	3,380	143,156
%	43.36	32.21	13.74	11.1	13.52	24.01	12.85	23.32
Secciones	654	1,863	247	293	425	1,922	151	5,555
%	40.12	88.28	11.72	8.67	11.27	21.03	12.36	20.57
Alumnos por región	36,713	154,543	45,400	63,072	78,972	208,906	26,313	613,919
% en total país	5.98	25.17	7.40	10.27	12.86	34.03	4.29	100.00
Cursos por región	1,630	5,758	2,107	3,379	3,772	9,138	1,222	27,006

Muestra de trabajo

Establecimiento	Ciudad de Bs. As.	Partidos GBA	Cuyo	Noreste	Noroeste	Pampeana	Patagonia	Totales
Público								
Alumnos	4,301	13,022	5,787	6,968	8,229	27,037	3,447	68,791
%	48.48	53.41	76.35	81.18	77.21	67.94	78.16	65.97
Secciones	525	1,682	654	854	996	2,967	401	8,079
%	54.74	59.25	80.54	84.55	80.00	71.70	81.01	70.26
Privado								
Alumnos	4,571	11,359	1,793	1,615	2,429	12,758	963	35,488
%	51.52	46.59	23.65	18.82	22.79	32.06	21.84	34.03
Secciones	434	1,157	158	156	249	1,171	94	3,419
%	45.26	40.75	19.46	15.45	20.00	28.30	18.99	29.74
Alumnos por región	8,872	24,381	7,580	8,583	10,658	39,795	4,410	104,279
% en total país	8.51	23.38	7.27	8.23	10.22	38.16	4.23	100.00
Cursos por región	959	2,839	812	1,010	1,245	4,138	495	11,498

Fuente: Elaboración propia en base a datos del ONE/2000.

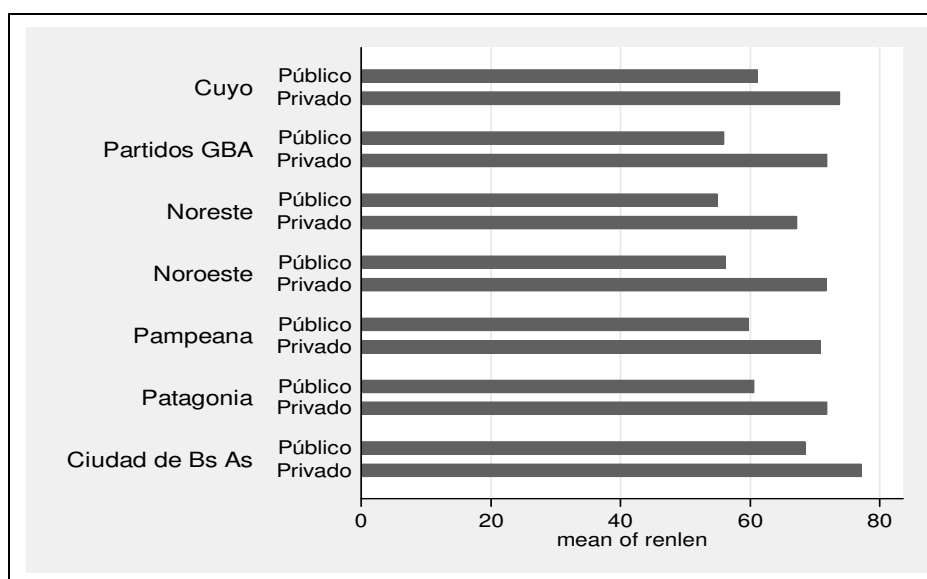
Comparando la primera línea en ambas secciones del cuadro se observa que la proporción de alumnos que concurren a establecimientos de gestión pública en la muestra, es menor que la correspondiente a la población, y ello para todas las regiones. Esto es debido a que la primera está compuesta por alumnos para los cuales se tiene información completa².

Esto plantea el interrogante de por qué los alumnos que asisten a escuelas privadas se muestran más capacitados para responder el cuestionario suministrado. Una posible explicación podría estar asociada con las notas.

Al respecto, el Gráfico I que muestra el nivel promedio de rendimientos por región para alumnos de escuelas públicas y privadas, puede darnos algunas claves sobre este sesgo hacia los privados de la muestra.

² La muestra de trabajo no es una muestra estrictamente estadística (en lo referente a su aleatoriedad), más bien es el resultado de considerar las observaciones para las cuales se tiene información en todas las variables incluidas en el estudio.

Gráfico I - Nivel promedio de Rendimiento en Lengua en las distintas regiones, desagregado por sector educativo



Fuente: Elaboración propia en base a datos del ONE/2000.

Como se observa en el Gráfico, en todas las regiones el rendimiento promedio en Lengua es más alto en las escuelas privadas que en las públicas, ilustrando la disparidad en el rendimiento promedio alcanzado por los distintos tipos de gestión educativa. Así, por ejemplo, el nivel alcanzado en el rendimiento escolar de un niño del Noroeste que concurre a un establecimiento público es bastante menor al rendimiento alcanzado por otro niño de la misma región que concurre a un establecimiento de gestión privada. Estas diferencias son menos notorias para los niños que residen en la Ciudad de Buenos Aires.

Este análisis llevaría a concluir que la muestra de trabajo contiene un problema de autoselección, ya que podría decirse que aquellos estudiantes que asisten a establecimientos privados estarían mejor preparados para responder las preguntas del cuestionario elaborado por el Ministerio de Cultura y Educación.

Sin embargo, las diferencias no son sustanciales, con lo que podemos esperar una buena representatividad de la población en la muestra.

II.3 Análisis de indicadores por regiones en la muestra de trabajo

Continuando con el análisis de la Tabla I, la última columna indica que el 66% de los alumnos en la muestra concurren a establecimientos de gestión pública. Por debajo de este valor se encuentran la Ciudad de Buenos Aires y los Partidos del Conurbano Bonaerense, mientras que el resto de las regiones se ubican por encima. El Noreste presenta el valor superior; con el 81% de sus alumnos asistiendo a escuelas estatales.

Por otra parte, en relación al total de establecimientos escolares, la misma tabla muestra que el 70% de las secciones son de gestión pública. La Ciudad de Buenos Aires y los Partidos del Conurbano Bonaerense presentan un porcentaje sensiblemente menor (55% y 60%, respectivamente).

La Tabla II resume las características de la muestra de trabajo que, siguiendo la metodología multinivel, se presentan ordenadas de acuerdo al nivel al que pertenecen las variables. El análisis se presenta para cada una de las regiones en que se divide al país.

La variable objeto de análisis (rendimiento en Lengua) presenta una gran disparidad entre la Ciudad de Buenos Aires y las restantes regiones. Mientras que los alumnos que pertenecen a la primera obtienen en promedio 75 puntos sobre 100 en la evaluación de Lengua,

estudiantes de las otras regiones obtienen entre 60 y 68 puntos promedio. Además, Ciudad de Buenos Aires tiene una población relativamente más homogénea, mientras que las restantes no sólo tienen un promedio más bajo sino además una población más heterogénea. Estos resultados indicarían que los niños que terminan sus estudios de nivel primario en la Ciudad de Buenos Aires poseen mayor preparación educativa que aquellos que pertenecen a cualquiera de las otras regiones³. Comparando con el nivel promedio general de 66 puntos, sólo Ciudad de Buenos Aires, Cuyo y Pampeana quedan por encima de este nivel. Sin embargo, la diferencia entre la primera y las dos restantes es de casi 8 puntos, estando más cerca, las últimas, de las otras regiones del país.

La proporción de alumnos repitentes en cada región también presenta una gran disparidad. Mientras que Ciudad de Buenos Aires muestra, junto con los Partidos del Conurbano Bonaerense, el menor porcentaje de alumnos con este atributo, con la menor variabilidad, regiones como el Noreste del país posee casi un cuarto de sus estudiantes como repitentes de algún grado, seguido muy de cerca por la Patagonia. Así, es mucho más probable que un chico sea repitente si asiste a una escuela que no pertenezca al Gran Buenos Aires, ya que en esta región alrededor de 1 de cada 12 alumnos es repitente, mientras que la proporción para el total del país es de 1 a 8.

Cuando se analiza el nivel socioeconómico de las regiones a través de un índice particular⁴, Noroeste es la región más desfavorecida y nuevamente Ciudad de Buenos Aires lidera el nivel alcanzado por todas las regiones. Esta última posee un nivel del índice de entre 11 y 17 puntos por encima de las restantes.

Muchos estudios incluyen al nivel socioeconómico como principal factor influyente en el rendimiento académico (Richard Rothstein, 2005; Cervini 2002b). El análisis del Gráfico II muestra la existencia de una marcada correlación entre el nivel medio del índice de nivel socioeconómico (INSE) y el promedio de las notas obtenidas en cada región. Así, la figura señala que mientras más elevada sea la media del nivel socioeconómico en una región, más alto es el promedio del rendimiento de Lengua⁵. Entre uno de los objetivos planteados en el trabajo se busca estudiar si esta relación es significativa y cuál es su magnitud.

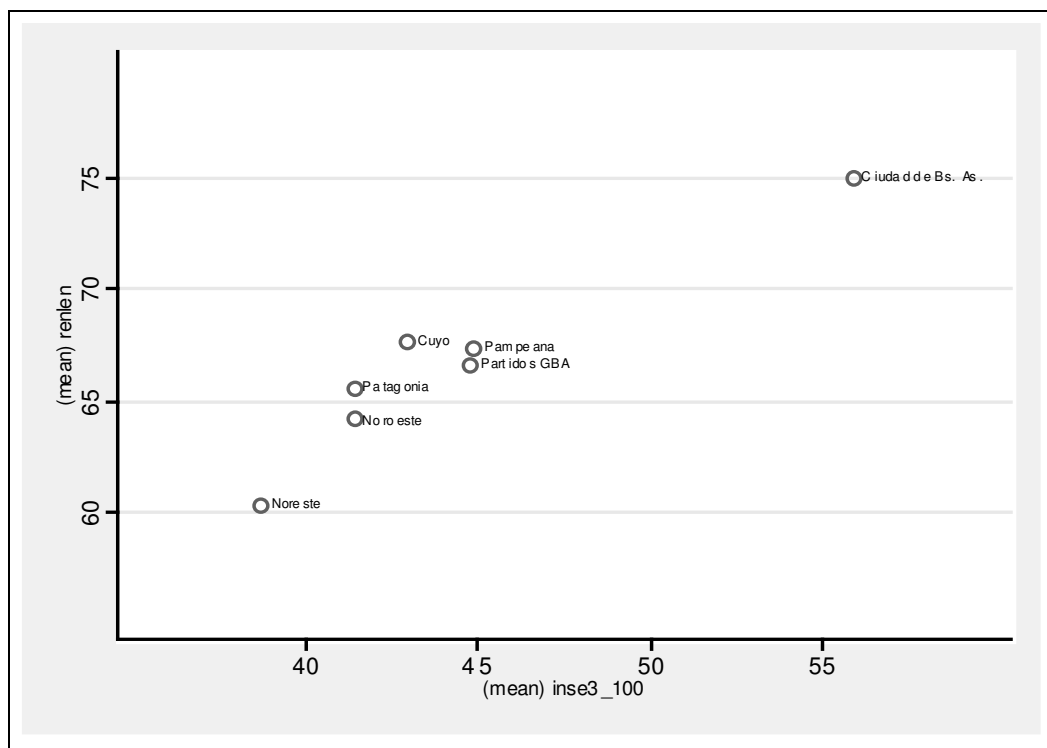
La comparación entre regiones nos lleva a ver la elevada distancia que separa a la región Noreste de la Ciudad de Buenos Aires. La gran disparidad existente en el nivel socioeconómico promedio de estas dos regiones coadyuva a acrecentar las diferencias en el rendimiento a través de otras variables de análisis, como es, por ejemplo, la deserción de hermanos y la concurrencia a diferentes sistemas educativos. Sin embargo, esta correlación estadística debiera estudiarse más detenidamente para averiguar los causales de esta asociación, como así también sobre otras variables que afectan al resultado de la prueba.

³ Existe una cuestión fundamental a tener en cuenta. Las pruebas son estandarizadas para todo el país y se prepara su contenido en la Ciudad de Buenos Aires, lo cual podría indicar que la metodología que se utiliza en la construcción de las pruebas está más en línea con la que se usa en la enseñanza de las escuelas de dicha región, sin tener en cuenta la forma en que se enseña en el resto del país. Esto podría crear una cierta desventaja para aquellos "buenos" alumnos que pertenecen a cualquier otra región. Sin embargo, este es un tema no desarrollado en el trabajo y que requiere una mayor profundización por parte de futuras investigaciones.

⁴La metodología de la construcción del INSE puede ser consultada en el Trabajo "Un índice de nivel socioeconómico de los hogares para aplicar a un modelo explicativo de rendimiento escolar", Verónica Herrero, Paula Palacios y Florencia Ruiz Díaz, 2005.

⁵ Lo que se afirma corresponde a un nivel de agregación de regiones y, por lo tanto, no implica que la misma relación se verifique cuando se consideran los alumnos como unidades de análisis, o en caso de darse, que no siempre sea de la misma magnitud la fuerza que vincula a ambas variables en las distintas regiones.

Gráfico II - Índice de Nivel Socioeconómico vs. Rendimiento Educativo en Lengua en las distintas regiones



Fuente: Elaboración propia en base a datos del ONE 2000.

Otras variables que se analizaron a nivel del alumno son la proporción de alumnos con hermanos desertores y el nivel educativo de las madres de los alumnos. Los valores arrojados muestran nuevamente a Ciudad de Buenos Aires liderando ambas variables. Esto también podría asociarse a lo que indica el nivel socioeconómico del hogar al que pertenece el estudiante, dado que los integrantes de hogares de muy bajo nivel socioeconómico poseen un alto costo de oportunidad al concurrir a la escuela, muchas veces quizás sin otra alternativa que ir a trabajar.

A nivel del curso, a su vez, se puede desagregar a las variables en aquellas relacionadas con el aula, con el docente y con el establecimiento. Estas variables caracterizan a los cursos, creando diferencias “entre” ellos.

El estado del aula hace referencia a la calidad en infraestructura con la que cuenta cada edificio educativo. El máximo nivel de este indicador lo alcanzan aquellas aulas que tienen “muy buena” iluminación, calefacción y ventilación, valoraciones a criterio de los docentes de cada sección. El Norte del país (regiones Noreste y Noroeste) presenta las aulas con menor calidad, quedando incluso muy por debajo del promedio general de 72,67 sobre 100.

Las características del docente a cargo de cada sección muestran a Ciudad de Buenos Aires con docentes de mayor años de experiencia en la profesión que cualquiera de las restantes regiones. Sin embargo, la región no lidera las variables de “Titularidad” y “Cursos de la Asignatura”. La mayor proporción de docentes titulares la posee la región Pampeana (76.2%), mientras que Cuyo tiene la mayor proporción de docentes que han tomado cursos relacionados a la asignatura que dictan (79,5%).

En el Apéndice I se describe como fueron construidas las distintas variables, así como el signo esperado.

Tabla II - Características de la muestra de trabajo

VARIABLES	Indicador	Ciudad de Buenos Aires	Partidos GBA	Cuyo	Noreste	Noroeste	Pampeana	Patagonia	Totales
Dependiente									
Rendimiento en Lengua (Índice, entre 0 y 100)	Media Desv. Estándar	75.030 16.655	66.611 19.136	67.666 17.871	60.407 19.098	64.192 18.859	67.414 18.541	65.583 17.947	66.909 18.806
Independientes									
Nivel del alumno									
<i>Propias del alumno</i>									
Sexo (Dummy, Varón =1)	Proporción Desv. Estándar	0.486 0.500	0.501 0.500	0.479 0.500	0.498 0.500	0.496 0.500	0.495 0.500	0.508 0.500	0.495 0.500
Repitencia (Dummy, Repitió algún grado =1)	Proporción Desv. Estándar	0.077 0.267	0.087 0.282	0.154 0.361	0.246 0.431	0.184 0.388	0.122 0.327	0.235 0.424	0.134 0.340
<i>Propias del hogar</i>									
Inse (Índice, entre 0 y 100)	Media Desv. Estándar	55.853 16.009	44.732 17.347	42.888 17.647	38.623 18.746	41.363 18.232	44.837 17.401	41.360 17.587	44.595 17.938
Deserción hermanos (Dummy, Hermanos que abandonaron la escuela =1)	Proporción Desv. Estándar	0.017 0.131	0.042 0.200	0.045 0.206	0.076 0.266	0.070 0.256	0.042 0.201	0.041 0.198	0.046 0.209
Libros en el Hogar (Dummy, Posee libros de la materia en el hogar= 1)	Proporción Desv. Estándar	0.831 0.375	0.848 0.359	0.810 0.393	0.799 0.401	0.828 0.377	0.837 0.369	0.831 0.375	0.833 0.373
Educación de la madre (Índice, entre 0 y 100)	Media Desv. Estándar	71.918 30.762	54.666 34.489	54.665 34.822	49.381 36.710	56.722 35.386	57.119 34.182	51.469 34.935	56.710 34.781
Nivel del curso									
<i>Propias del aula</i>									
Estado del aula (Índice, entre 0 y 100)	Media Desv. Estándar	82.360 19.988	71.504 25.635	77.514 23.742	58.244 28.106	60.871 26.685	75.824 24.229	79.577 21.505	72.676 25.635
<i>Propias del docente</i>									
Educación del docente (Años de escolaridad)	Media Desv. Estándar	14.957 0.824	14.774 0.655	14.928 0.573	14.990 0.457	14.907 0.601	14.842 0.649	14.865 0.606	14.862 0.646
Experiencia del docente (Años de antigüedad)	Media Desv. Estándar	16.434 9.291	12.018 8.091	12.594 7.451	12.728 7.548	13.741 8.060	13.600 7.856	11.674 7.241	13.259 8.076
Titularidad (Dummy, Es titular =1)	Proporción Desv. Estándar	0.749 0.434	0.750 0.433	0.684 0.465	0.702 0.457	0.739 0.439	0.762 0.426	0.586 0.493	0.738 0.440
Cursos de la asignatura (Dummy, Tomó cursos de la asignatura = 1)	Proporción Desv. Estándar	0.728 0.445	0.502 0.500	0.795 0.404	0.662 0.473	0.773 0.419	0.710 0.454	0.729 0.445	0.672 0.469
<i>Propias de la escuela</i>									
Sector (Dummy, Sector de gestión privado = 1)	Proporción Desv. Estándar	0.515 0.500	0.466 0.499	0.237 0.425	0.188 0.391	0.228 0.420	0.321 0.467	0.218 0.413	0.340 0.474
Interacciones									
Inse por Sector (Índice, entre 0 y 100)	Proporción Desv. Estándar	61.892 13.705	51.869 16.747	56.421 15.734	53.448 17.251	56.274 15.582	53.623 16.184	52.754 16.687	54.418 16.382
Deserción por repitencia (Dummy, si repitió algún grado y tiene hermanos desertores = 1)	Proporción Desv. Estándar	0.060 0.237	0.149 0.356	0.116 0.320	0.142 0.349	0.153 0.360	0.126 0.332	0.078 0.269	0.091 0.287
Total de observaciones		8,872	24,381	7,580	8,583	10,658	39,795	4,410	104,279

Fuente: Elaboración propia en base a datos del Operativo Nacional de Evaluación de la Calidad Educativa del año 2000 (ONE/2000).

III. Modelo estadístico

Como se anticipó en la introducción los estudios empíricos de funciones de producción más actuales se apoyan en el desarrollo de técnicas que reconocen los distintos "niveles" en los que los "insumos educativos" pueden influir en la creación de conocimiento (Rowe K., Hill P., Holmes-Smith P., 1995; Cervini 1999, 2002 y 2004; Herrera Gómez y otros, 2005).

En esta sección se resumen las características de los modelos multinivel, y en la siguiente Sección IV se presentan los resultados de las estimaciones para cada región. Un primer modelo de componente de varianza, frecuentemente denominado ANOVA, que permite determinar la proporción de la varianza total de la variable a explicar, en este caso, el rendimiento en Lengua, imputable a diferencias en los cursos. En el mismo no se incorporan variables explicativas. El segundo análisis consiste en la estimación de un modelo multinivel. En este se incluyen las variables explicativas en dos niveles, a saber, nivel del alumno (que puede considerarse el nivel más "bajo") y nivel del curso o clase (nivel más "alto").

III.1 Modelo de componente de varianzas o ANOVA

Este modelo tiene la finalidad de descomponer la varianza en el rendimiento en la prueba de Lengua en dos partes, atendiendo a la estructura jerárquica de los datos. Así, puede determinarse, de la variabilidad total en el puntaje de los alumnos en Lengua, la proporción que puede ser atribuido al efecto de los cursos y la que puede ser explicada por las variables propias del alumno. En este sentido, por ejemplo, si el modelo indica que una proporción del 30% de la variabilidad en las notas es explicada por el curso, ello puede interpretarse como que sólo un máximo del 70% de la variabilidad puede atribuirse a las características propias del alumno y su entorno. En resumen, el ANOVA constituye el primer paso que justifica la estimación de un modelo multinivel en el que se incorporen variables explicativas en distintos niveles.

Esquemáticamente, lo que realiza el modelo ANOVA es lo siguiente:

$$\text{Variabilidad Notas} = \alpha * \text{Variabilidad dentro del curso} + (1-\alpha) * \text{Variabilidad entre cursos} + \text{Error}$$

Trabajando sobre la fórmula anterior es posible obtener un coeficiente llamado "Rho" que indica la proporción de la variabilidad total del rendimiento en Lengua que puede atribuirse a las características del curso. Un "Rho" más alto arroja indicios de una estructura jerarquizada de los datos, en el sentido de que una mayor proporción del movimiento del rendimiento en Lengua de los alumnos respecto a la media general se concentra en el nivel del curso. En otras palabras, los cursos aparecen como "estructuras homogéneas" en el sentido de que existe una influencia mutua entre los alumnos pertenecientes a los mismos.

III.2 Modelo Multinivel

El modelo multinivel incorpora las variables explicativas del nivel del alumno (nivel "más bajo") y del curso (nivel "más alto"). El propósito es intentar asociar las variables específicas de cada nivel que contribuyen a explicar la importancia relativa que adquiere cada uno de los niveles en la variabilidad total. Siguiendo con el ejemplo anterior, se esperaría que el 30% de la variabilidad total que ha sido atribuida al aula resulte explicada por la experiencia de los docentes, calidad edilicia, etc.; mientras que el 70% atribuido al

nivel individual del alumno refleje el efecto de las características propias del mismo (sexo, repitencia, etc.) y de su entorno familiar. El desarrollo matemático del modelo se presenta en el Apéndice II del trabajo. Este modelo se aplicó a los datos presentados en la sección anterior para cada una de las regiones del país, y cuyos resultados se presentan en la sección siguiente.

IV. Resultados obtenidos

Para comenzar, aplicamos el modelo de análisis de la varianza para descomponer la variabilidad de las notas en sus dos componentes, a saber, la variabilidad atribuida a las diferencias entre los alumnos, y la correspondiente a las diferencias entre las aulas. Este modelo se estima para cada una de las regiones a comparar. Los resultados obtenidos se presentan en la Tabla III siguiente, cuya última fila indica el coeficiente “Rho” que resume la información acerca de la importancia relativa de la escuela dentro de la variabilidad total. Los valores obtenidos para “Rho” sugieren que el agrupamiento explica alrededor de un 30% del rendimiento alcanzado por cada alumno. Las regiones de menor desarrollo relativo (Noreste y Noroeste) presentan un valor más alto para el “Rho”, mientras que el coeficiente más bajo se observó para la Ciudad de Buenos Aires, seguido muy de cerca por el de la Patagonia. Como resultado de lo anterior creemos que es necesario la utilización de un análisis multinivel, incorporando el efecto anidamiento de los niños en las escuelas.

Tabla III- Descomposición de varianza

VARIABLES	Cuyo	Partidos GBA	Ciudad de Buenos Aires	Noreste	Noroeste	Pampeana	Patagonia
Intercepto σ_{00}^2	101.83	134.24	70.83	138.47	138.62	105.34	85.61
Nivel alumno σ_e^2	219.22	238.33	211.62	222.50	221.94	243.20	239.69
Rho	0.32	0.36	0.25	0.38	0.38	0.30	0.26

Fuente: Elaboración propia en base a datos del Operativo Nacional de Evaluación de la Calidad Educativa del año 2000 (ONE/2000).

Los resultados referidos al modelo multinivel se presentan en la Tabla IV. Para cada región, la tabla muestra los coeficientes asociados con el efecto de cada una de las variables descritas en la primera columna sobre las notas.

Considerando estos resultados podemos observar que existen diferencias importantes entre regiones en el comportamiento de las distintas variables del modelo. En particular, las regiones de menor riqueza relativa presentan un mayor peso de las variables asociadas a la escuela, mientras que en las de mayor riqueza ocurre lo contrario. Así, esto sugiere que el trabajo realizado dentro del aula impacta con mayor fuerza sobre los resultados de los alumnos que residen en las regiones más pobres. Por lo tanto, parecería que la escuela, como vehículo de promoción social, forma una parte importante de las políticas de igualación de oportunidades en las regiones pobres.

Con respecto a las variables propias del alumno, expuestas en las dos primeras filas de la Tabla IV, se observa que los varones obtienen en relación a las mujeres una nota

promedio inferior en 5 puntos, fenómeno que se observa con ligeras variaciones en todas las regiones. Haber repetido grado tiene efecto de entre -8.5 y -10 puntos en las regiones ricas (Ciudad de Buenos Aires, Partidos de GBA, Cuyo, Patagonia y Pampeana), mientras que en las regiones de menor desarrollo relativo el efecto es sensiblemente menor (alrededor de -7 puntos para Noreste y Noroeste). Esto estaría indicando una mayor penalización a la condición de repitente en las regiones más ricas.

Con relación a la condición socioeconómica del hogar, se analizaron 4 indicadores: el INSE, tener hermanos desertores, nivel educativo de la madre, y la tenencia de libros específicos de Lengua en el hogar, siendo las dos primeras las relevantes para este trabajo. El indicador del nivel socioeconómico (INSE), muestra marcadas diferencias entre las regiones según su grado de desarrollo. Así, la Ciudad de Buenos Aires presenta efecto positivo y elevado (1.2 puntos por decil de ingreso), mientras que en las regiones donde la pobreza pesa mucho en la condición socioeconómica promedio (NEA, Partidos GBA y NOA) el efecto es casi nulo. Por otro lado, la presencia de hermanos desertores tiene un efecto similar a la repitencia del propio alumno, aunque de una magnitud mayor. Por ejemplo, un alumno de Ciudad de Buenos Aires que tenga algún hermano que haya abandonado la escuela obtiene 15 puntos menos en las pruebas; por el contrario, un estudiante con las mismas características en otras regiones del país obtiene entre 8 y 10 puntos menos.

Esto confirmaría algunos resultados expuesto en la literatura comparada donde se señala que en los países más industrializados el principal determinante de los logros educativos está asociado con el hogar, mientras que en los países de menor desarrollo relativo, la escuela juega un papel más importante en la determinación de dichos logros (Coleman et al, 2000; Lockheed y Longford, 1989).

Con respecto a las variables asociadas al curso, se estudió el estado edilicio del aula. En general se observa que una mayor calidad del aula está asociada a una mayor nota de los alumnos, aunque no presenta una magnitud importante en las diferentes regiones.

Las variables del docente que se analizaron fueron 4: la educación, la titularidad, la realización de cursos en la asignatura y la experiencia del docente. En cuanto a la educación, tiene un efecto positivo solo en 3 regiones (Noreste, Pampeana y Partidos de GBA), mientras que en las restantes no es muy diferente de cero. Esto implicaría que un docente con más años de educación no ejerce una influencia importante sobre el puntaje obtenido en la prueba. Los cursos específicos en Lengua tomados por los docentes no afectan al rendimiento de los alumnos en las regiones de Noroeste, Noreste, Ciudad de Buenos Aires y Patagonia. La experiencia del docente tiene un efecto positivo sobre el resultado de las pruebas en todas las regiones a excepción de la Ciudad de Buenos Aires, donde no se puede concluir que el efecto sea distinto de cero. Esta asociación positiva entre el rendimiento en la prueba y la experiencia se presenta con más fuerza en las regiones con menor desarrollo relativo.

Finalmente, el sector de gestión del establecimiento al que asisten los alumnos tiene un efecto importante sobre el rendimiento en Lengua en todas las regiones, si bien existen diferencias en la magnitud del mismo. Por ejemplo, los alumnos que asisten a escuelas privadas en regiones con mayor desarrollo económico tienen entre 5 y 7 puntos más de nota que el promedio, mientras que en las regiones más pobres este efecto está en el orden de los 9 a 10 puntos.

Tabla IV – Resultados Análisis Multinivel

VARIABLES	Cuyo	Partidos GBA	Ciudad de Buenos Aires	Noreste	Noroeste	Pampeana	Patagonia
Nivel del alumno							
<i>Propias del alumno</i>							
Sexo (Dummy, Varón =1)	-5.2576 0.0000	-5.6269 0.0000	-5.1441 0.0000	-5.1520 0.0000	-4.7521 0.0000	-6.3765 0.0000	-5.9444 0.0000
Repitencia (Dummy, Repitió algún grado =1)	-8.5464 0.0000	-9.7388 0.0000	-10.3698 0.0000	-6.7051 0.0000	-6.8223 0.0000	-9.3362 0.0000	-8.2821 0.0000
<i>Propias del hogar</i>							
Inse (Índice, entre 0 y 100)	0.0658 0.0000	0.0169 0.1250	0.1248 0.0000	0.0249 0.0980	0.0141 0.3130	0.0906 0.0000	0.0602 0.0060
Deserción hermanos (Dummy, Hermanos que abandonaron la escuela =1)	-9.2233 0.0000	-9.9878 0.0000	-14.8570 0.0000	-8.0814 0.0000	-8.7597 0.0000	-10.5369 0.0000	-10.7216 0.0000
Educación de la madre (Índice, entre 0 y 100)	0.0188 0.0070	-0.0053 0.1800	0.0041 0.5590	0.0277 0.0000	0.0158 0.0080	0.0124 0.0000	0.0204 0.0310
Libros en el Hogar (Dummy, Posee libros de la materia en el hogar= 1)	0.8069 0.0770	0.9068 0.0010	0.3041 0.4720	0.9325 0.0320	0.9213 0.0240	1.6453 0.0000	1.1119 0.0810
Nivel del curso							
<i>Propias del aula</i>							
Estado del aula (Índice, entre 0 y 100)	0.0767 0.0000	0.0857 0.0000	0.0445 0.0000	0.0212 0.1050	0.0282 0.0190	0.0509 0.0000	0.0404 0.0230
<i>Propias del docente</i>							
Educación del docente (Años de escolaridad)	0.1814 0.7540	0.8624 0.0030	-0.0930 0.7660	1.4545 0.0710	0.0421 0.9400	0.4198 0.0720	0.4488 0.4740
Titularidad (Dummy, Es titular =1)	-0.8693 0.2490	-0.9349 0.0440	0.3555 0.5650	-0.4763 0.6190	1.8824 0.0170	0.2807 0.4890	-0.3926 0.6490
Cursos de la asignatura (Dummy, Tomó cursos de la asignatura = 1)	1.5399 0.0400	0.5468 0.1350	0.5389 0.3340	0.2025 0.7860	1.1467 0.1130	1.5097 0.0000	-0.0321 0.9700
Experiencia del docente (Años de antigüedad)	0.1687 0.0010	0.1479 0.0000	0.0057 0.8510	0.2224 0.0000	0.1489 0.0010	0.0736 0.0010	0.1358 0.0200
<i>Propias de la escuela</i>							
Sector (Dummy, Sector de gestión privado = 1)	7.1048 0.0000	10.6516 0.0000	5.1619 0.0000	9.0376 0.0000	9.0134 0.0000	5.7052 0.0000	7.0045 0.0000
Interacciones							
Inse por Sector (Índice, entre 0 y 100)	0.0797 0.0060	0.0652 0.0000	0.0044 0.8490	0.0340 0.2340	0.0806 0.0020	0.0275 0.0190	0.0553 0.1420
Deserción por repitencia (Dummy, si repitió algún grado y tiene hermanos desertores = 1)	2.8195 0.1000	6.9218 0.0000	10.5762 0.0000	3.5593 0.0040	4.8330 0.0000	6.6124 0.0000	5.2329 0.0270
Constante	62.3480 0.0000	50.3769 0.0000	75.6299 0.0000	36.2048 0.0030	58.6996 0.0000	59.0231 0.0000	58.8056 0.0000
Varianza							
Intercepto σ_0^2	51.83	64.49	34.71	96.73	86.58	63.42	43.08
Nivel alumno σ_e^2	201.24	222.18	197.00	205.70	208.83	221.03	216.67
N° de Observaciones							
Alumnos	7580	24381	8872	8583	10658	39795	4410
Secciones	812	2839	959	1010	1245	4138	495

Fuente: Elaboración propia en base a datos del Operativo Nacional de Evaluación de la Calidad Educativa del año 2000 (ONE/2000).

V. Síntesis de conclusiones

En el trabajo se intentó determinar la fuerza de la asociación entre el rendimiento obtenido en las pruebas estandarizadas en Lengua y un conjunto de variables indicativas de las características del alumno y su entorno familiar, del aula, del docente y de la escuela a la que asiste. Se analizó la existencia de posibles diferencias regionales en la sensibilidad de estas asociaciones. Con tal fin, se describieron los principales indicadores, se estimó un modelo multinivel en dos etapas para cada una de las regiones en que se dividió el país, y luego se compararon los resultados obtenidos.

El estudio descriptivo se apoyó en la información que brindó el Operativo Nacional de Evaluación del año 2000 (ONE/2000). El mismo indicó entre las regiones existe una gran disparidad en lo que respecta al hogar de los estudiantes, por un lado, y a las características del docente y la escuela, por el otro. Por ejemplo, el INSE correspondiente a la Ciudad de Buenos Aires presentó valores que reflejan una situación socioeconómica del hogar ampliamente superior en esta región, en relación a las restantes. Por otra parte, también la Ciudad de Buenos Aires resultó ser la región con menor proporción de alumnos repitentes en todo el país, mientras que en las regiones de menor desarrollo relativo se observó la mayor proporción de alumnos con estas características.

Con respecto al nivel de los efectos de las variables contempladas, su análisis por regiones indicó la existencia de sensibles diferencias de nivel para la mayoría de las variables analizadas.

Más específicamente, el estudio puso de manifiesto que las variables del nivel “más bajo”, asociadas con la variabilidad entre alumnos (sexo, condición de repitente, INSE, hermanos desertores, etc.), en conjunto, presentaron mayor fuerza explicativa en las regiones de mayor desarrollo relativo, mientras que, por el contrario, las variables del nivel “más alto” (educación del docente, experiencia, sector, etc.) aumentaron su valor explicativo significativamente en las regiones más pobres.

En suma, el trabajo sugiere la conveniencia de llevar a cabo una política de igualación de oportunidades educativas regionalizada, de manera de poder focalizar los recursos de forma más efectiva, más apropiada a cada región.

VI. Referencias Bibliográficas

- Brik A, Raudenbush, Stephen W. (1992). *Hierarchical Linear Models: Applications and Data Analysis Methods*. Sage Publications, California.
- Brik A, Raudenbush, Stephen W.; Seltzer, M.; Congdon, Richard T. (1986). *An Introduction to HLM: Computer Program and User's Guide*. University of Chicago, Chicago.
- Cervini R (2002b). Desigualdades Socioculturales en el Aprendizaje de Matemática y Lengua de la Educación Secundaria en Argentina: Un Modelo de Tres Niveles. *RELIEVE: Revista Electrónica de Investigación y Evaluación Educativa* **V. 8**, Pp. 135-158.
- Cervini R (2004a). Influencia de los Factores Institucionales sobre el Logro en Matemática de los Estudiantes en el Último Año de la Educación Media de Argentina - Un Modelo de tres Niveles. *REICE - Revista Electrónica Iberoamericana sobre Calidad, Eficacia y Cambio en Educación* **V. 2**.
- Cervini R (2004b). Nivel y Variación de la Equidad en la Educación Media de Argentina. *Revista Iberoamericana de Educación (ISSN: 1681-5653)*.
- Goldstein H (1991). Multilevel Modelling of Survey Data. *The Statistician* **V. 40**, Pp. 235-244.
- Goldstein H (1995). *Multilevel Statistical Models* 2nd ed. New York Halstead Press.
- Herrera Gomez M, Aráoz, M. Florencia, de Lafuente, Gisela, D`jorge, M. Lucrecia, Granado, M. Jose, Michel Rivero, Andrés, Paz Terán, Corina (2005). Técnicas para Datos Multinivel: Aplicación a los Determinantes del Rendimiento Educativo. In *Anales de la XL Reunión Anual de la Asociación Argentina de Economía Política*, La Plata.
- Hox JJ (1995). *Applied Multilevel Analysis*, pp. 126. TT-Publikaties.
- Levin HM (1971). A New Model of School Effectiveness, pp. 55-75. Secretary of HEW, Washington D.C.
- Lockheed M, Longford N (1989). A Multilevel Model of School Effectiveness in a Developing Country. *World Bank Discussion Papers*. **Nº 69**.
- Morduchowicz A (2003). *Discusiones de Economía de la Educación*. Ed. Losada. Buenos Aires.
- Rothstein R (2004). A Wider lens on the Black-White Achievement Gap. *Phi Delta Kappan*. **V. 86 Nº 2**, pp.104-110.
- Rowe KJ, Hill, Peter W., Holmes-Smith, Philip (1995). Methodological Issues in educational Performance and School Effectiveness Research: A Discussion with Worked Examples. *Australian Journal of Education* **V. 39 Nº 3**, pp. 217-248.
- Singer JD (1998). Using SAS PROC MIXED to Fit Multilevel Models, Hierarchical Models, and Individual Growth Models. *Journal of Educational and Behavioral Statistics* **V. 24**, Pp. 323-355.

Apéndice I

Tabla de definiciones y construcción de variables

VARIABLES	SIGNO ESPERADO	TIPO	CONSTRUCCION ^(a) partir del cuestionario del alumno y el docente de las bases ONE 2000)	REGRESIÓN
Nivel del alumno				
<i>Propias del alumno</i>				
Sexo	(-)	Dummy	Se asignó el valor 0 para las mujeres y 1 para los varones.	
Repitencia	(-)	Dummy	Se asignó el valor 0 a los alumnos que nunca repitieron grado y 1 a los que alguna vez lo hicieron.	
<i>Propias del hogar</i>				
Inse	(+)	0-100 (continua)	Se tuvieron en cuenta tres variables "Nivel de instrucción del padre", "Posesión de 18 bienes en el hogar", "Condición de hacinamiento". (*)	Variable centrada respecto a su media
Deserción hermanos	(-)	Dummy	Se asignó el valor 1 si el alumno tiene hermanos entre 7 y 17 años que nunca fue o que abandonó la escuela, y 0 si no los tiene.	
Libros en el Hogar	(+)	Dummy	Se asignó el valor 1 si el alumno posee algún libro de matemática o lengua, según el caso, en el hogar y 0 en el caso que no posea libro alguno.	
Educación de la madre	(+)	0-100 (6 niveles)	A partir de la preg. 9 se recodificó con 0 (primario incompleto), 1 (primario completo), 2 (secundario incompleto), , 5 (terciario o universitario completo). Luego se llevó el rango de 0 a 100.	
Nivel del curso				
<i>Propias del aula</i>				
Estado del aula	(+)	0-100 (10 niveles)	A partir de la preg. 31 del cuestionario del docente relativa a la opinión de "iluminación", "calefacción" y "ventilación", se realizó un índice mediante el cociente entre la suma de las respuestas de los tres indicadores y la cantidad de respuestas sumadas en el numerador. El rango fue llevado de 0 a 100.	Variable centrada respecto a su media
<i>Propias del docente</i>				
Educación del docente	(+)	(12-17)	Se generó a partir de la preg. 6 del cuestionario del docente "¿Qué títulos ha obtenido Ud.?" y se le asignó un valor en número de años estudiados con un mínimo de 12 (si no tiene título) hasta un máximo de 17 (título de Profesional Universitario).	
Experiencia del docente	(+)	0.5-35 (8 niveles)	Se refiere a la antigüedad en la profesión de docente. Se asignó la marca de clase a intervalos construídos en años que van de "menos de un año" a "más de 30 años" con saltos cada 5 años.	
Titularidad	(+)	Dummy	Se asignó el valor 1 si el docente era titular y 0 si no lo era.	
Cursos de la asignatura	(+)	Dummy	Se asignó el valor 1 si el docente realizó cursos específicos de la materia que dicta y 0 en caso contrario.	
<i>Propias de la escuela</i>				
Sector	(+)	Dummy	Se asignó el valor 1 si el establecimiento al que concurre el alumno era de gestión privada y 0 si era de gestión pública.	
Interacciones				
Inse por sector	?		Producto de variables indicadas	
Deserción por repitencia	El efecto se potencia		Producto de variables indicadas	

Notas: (*) La metodología para la construcción del índice puede ser consultada en el paper "Un índice de nivel socio-económico de los hogares para aplicar a un modelo explicativo del rendimiento escolar"

Apéndice II

Desarrollo estadístico de los modelos ANOVA y multinivel

-Modelo de Componente de varianzas o ANOVA

Siguiendo a Hox (1995) y Singer (1998), tomando el subíndice i para referirse al alumno y j para referirse al curso, el modelo estadístico de componente de varianza puede describirse a través de las siguientes ecuaciones:

$$Y_{ij} = \beta_{0j} + \varepsilon_{ij} \quad (1)$$

$$\beta_{0j} = \gamma_{00} + v_{0j} \quad (2)$$

En la ecuación (1) Y_{ij} representa el rendimiento en Lengua del alumno i en el curso j ; el mismo se expresa como la suma de una ordenada al origen β_{0j} , que se admite puede variar entre aulas, y un término de error ε_{ij} para el individuo i en el curso j , que se asume tiene una distribución $\varepsilon_{ij} \sim N(0, \sigma^2_{\varepsilon})$.

En (2) se descompone la ordenada al origen del curso j en γ_{00} (promedio del rendimiento en Lengua para el total de alumnos) y un término de error v_{0j} que se presenta al nivel de la clase. Con respecto a este último, también se supone que $v_{0j} \sim N(0, \sigma^2_{00})$.

Combinando (1) y (2) se obtiene:

$$Y_{ij} = \gamma_{00} + v_{0j} + \varepsilon_{ij} \quad (3)$$

El modelo de la ecuación (3) no contiene variable explicativa alguna; sólo descompone la nota en Lengua en γ_{00} , v_{0j} y ε_{ij} . Se admite que el rendimiento en Lengua de un alumno puede diferir del promedio general (γ_{00}) debido a las características propias del curso, englobadas en v_{0j} , y las propias de él mismo y su entorno familiar, representadas por ε_{ij} .

A partir de (3) se puede obtener estimaciones de las varianzas σ^2_{ε} y σ^2_{00} necesarias para el cálculo del coeficiente *intra-clase* ("Rho"), que se define como:

$$\rho = \frac{\sigma^2_{00}}{\sigma^2_{\varepsilon} + \sigma^2_{00}}$$

Por tanto, ρ no es otra cosa que la proporción de la variabilidad total del rendimiento en Lengua que puede atribuirse a las características del curso. Un ρ más alto arroja indicios de una estructura jerarquizada de los datos en el sentido de que una mayor proporción del movimiento del rendimiento en Lengua de los alumnos respecto a la media general se concentra en el nivel del curso.

-Modelo Multinivel

En éste se incorporan variables explicativas del nivel del alumno (nivel "más bajo") y del nivel del aula (nivel "más alto"). El objetivo es determinar la significación y magnitud de los

efectos de estas variables y analizar la porción adicional de varianza explicada por el modelo.

Para describirlo, se sigue la estrategia de Bryk y Raudenbush (1986 y 1992) de modelar los coeficientes no observados al nivel del alumno en el nivel del curso. En este caso, el modelo puede ser descrito convenientemente por las siguientes ecuaciones:

$$Y_{ij} = \beta_{0j} + \beta_p * X_{p_{ij}} + \varepsilon_{ij} \quad (4)$$

$$\beta_{0j} = \gamma_{00} + \gamma_{0k} * Z_{kj} + \nu_{0j} + \varepsilon_{ij} \quad (5)$$

Combinando las ecuaciones (4) y (5) según la sugerencia de Goldstein (1995) se tiene:

$$Y_{ij} = \gamma_{00} + \beta_p * X_{p_{ij}} + \gamma_{0k} * Z_{kj} + \nu_{0j} + \varepsilon_{ij} \quad (6)$$

En (6) se expresa el rendimiento en Lengua como resultado de una combinación de factores: las variables del nivel del alumno que describen sus características personales y las de su entorno familiar, representadas con X_{p_i} ; las variables que caracterizan al curso, denominadas genéricamente con Z_k ; y los términos de error que cumplen con los mismos supuestos arriba mencionados. Esta ecuación puede considerarse como la función de producción de educación.

Las variables X_p incorporadas en el presente trabajo son el género o sexo del alumno, la repitencia, el índice de nivel socioeconómico (INSE), la deserción de los hermanos, la posesión de libros en el hogar y el nivel educativo de la madre. Las variables Z_k del nivel "más alto" incluidas son las del curso propiamente dichas, como el estado del aula; las que se relacionan con el docente, como la educación, la titularidad, cursos realizados sobre la asignatura y la experiencia del mismo en la profesión. Además, se incluye una variable sector que permite caracterizar las escuelas a las que pertenecen los cursos según sean de gestión estatal o privada. Finalmente, se agregan también dos interacciones entre las variables sector e índice socioeconómico y deserción y repitencia.