



DOCUMENTOS DE TRABAJO

Del PROYECTO PICT 2003 N° 14534

FONCYT / SECYT

DOCUMENTO DE TRABAJO N° 5 / 2006

**Análisis multinivel del rendimiento escolar
al término de la educación básica en
Argentina**

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

Noviembre 2006

INSTITUTO DE ECONOMÍA Y FINANZAS

FACULTAD DE CIENCIAS ECONÓMICAS

UNIVERSIDAD NACIONAL DE CÓRDOBA

**Análisis multinivel del rendimiento escolar al término de la
educación básica en Argentina**

**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
Agosto de 2006**

Para ser presentado en la Reunión Anual de la Asociación Argentina de Economía Política,
Salta, 8 al 10 de Noviembre de 2006

Análisis multinivel del rendimiento escolar al término de la educación básica en Argentina*.

Héctor Gertel, Roberto Giuliadori, Verónica Herrero,
María Luz Vera, Diego Fresoli, Guadalupe Morra
Facultad de Ciencias Económicas
Universidad Nacional de Córdoba
hgertel@eco.unc.edu.ar

1. INTRODUCCIÓN

El estudio de los factores que contribuyen a explicar el rendimiento escolar ejerce en la actualidad una considerable atracción entre los economistas, porque mejorar la eficiencia de las escuelas se ha convertido en un objetivo central de la acción pública en muchos países preocupados por incrementar la contribución de la inversión educativa a su desarrollo económico y social.

El trabajo tiene como propósito estudiar los factores que contribuyen al rendimiento escolar en la Argentina al término de la educación básica, aplicando técnicas de análisis multinivel.

Cuando las observaciones, como en este caso de la función de producción de educación, aparecen agrupadas y jerarquizadas en niveles: alumnos en cursos, cursos en escuelas y escuelas en jurisdicciones la aplicación de técnicas de análisis lineal jerárquico de datos, o multinivel arroja, en relación al análisis de regresión lineal multivariable convencional resultados más potentes.

En estos casos, los modelos de multinivel facilitan la interpretación de la variabilidad observada en los resultados alcanzados en las pruebas escolares por un conjunto de i alumnos, agrupados en j aulas de k escuelas localizadas en diferentes jurisdicciones. En general, se espera que la variabilidad de las notas que miden el resultado del aprendizaje de los alumnos agrupados dentro de un mismo curso presente un cierto grado de homogeneidad, mientras que la mayor variabilidad se encontraría entre cursos, escuelas y jurisdicciones. Este enfoque, en suma, contribuye a identificar aquella porción de la variabilidad total en las notas específicamente asociadas con el aporte de las escuelas, a diferencia de las funciones de producción convencionales donde las variables del alumno, de su hogar y de la escuela asumen un mismo nivel de jerarquía. Ayuda a identificar a las escuelas eficaces, diferenciándolas de las que no lo son tanto y permite medir la contribución individual de los recursos (docentes, equipamiento, sistemas de gestión) a la eficacia de las escuelas.

El trabajo está organizado de la siguiente manera: la sección siguiente desarrolla de manera sencilla el modelo conceptual que se propone aplicar. La sección tres describe los datos, señalándose las principales características de los mismos. La sección cuatro contiene el desarrollo del modelo a estimar mientras que en la sección cinco se presentan los resultados principales de las estimaciones realizadas y se comparan con los obtenidos por MCO. Una sección final es reservada para discutir y resumir el alcance de los resultados que arroja este estudio.

* Los autores agradecen los comentarios recibidos a una presentación previa del trabajo realizada en las I Jornadas Hispano – Argentinas de Economía Pública, en Badajoz, España, 06-07 de Julio de 2006. El financiamiento fue provisto por el Proyecto PICT-03 N° 14534.

2. EL MODELO CONCEPTUAL

El modelo conceptual se apoya en la noción de función de producción educativa, que representa el punto de partida para estudiar la relación de efectividad existente entre el resultado del aprendizaje en las escuelas y el conjunto de recursos disponibles con que éstas cuentan para resolver el proceso de enseñanza y aprendizaje (Levin, 1983, Cooms y Hallak, 1987).

El proceso de formación de una persona para asumir su papel en la sociedad es complejo y no solo atañe a los aprendizajes impartidos en la escuela. También importan los aprendizajes tomados del hogar, del medio y de las etapas escolares anteriores. El efecto unitario que produce sobre el aprendizaje cada uno de los factores que intervienen en la escuela, el hogar y el medio está, a su vez, influenciado por las características personales de los alumnos, que se expresan en diferencias de habilidad, motivación y predisposición hacia el aprendizaje. Coleman y sus colaboradores (Coleman, et.al. 1966) en uno de los primeros esfuerzos por detectar relaciones estadísticas entre el aprendizaje, el hogar y la escuela utilizaron un método simple de agregación de variables y concluyeron que el hogar ejercía, en los Estados Unidos, la principal influencia sobre los resultados de las pruebas escolares y que la escuela importaba poco. Bowles y Levin (1968), demolieron las conclusiones del Informe Coleman demostrando que, en ese estudio, las técnicas de "incorporación progresiva de variables" habían sido aplicadas sin advertir la existencia de una alta correlación entre las variables representativas del hogar y de la escuela que invalidaba los resultados obtenidos (si se hubiera invertido el orden de incorporación se habría obtenido la conclusión inversa: solo la escuela cuenta y el hogar de pertenencia no afecta a los resultados). En el mismo estudio los autores especifican una función de producción de tipo lineal multivariada para re-estimar, con los mismos datos que utilizó Coleman, el efecto escuela, el efecto hogar y el efecto conjunto, mostrando que al corregir la metodología se obtiene que la escuela posee, en Estados Unidos, un efecto positivo y significativo sobre el aprendizaje, si éste es medido por el puntaje obtenido en pruebas de rendimiento escolar. Levin (1970), también partiendo de una especificación lineal de la función de producción escolar, produjo cálculos indicativos de una relación de efectividad a costos para la contratación de maestros jóvenes, con título universitario, que resultó superior a la asociada con la contribución al rendimiento escolar de incrementar el valor de los premios por experiencia, si se utilizaba una escala de premios típica. En un estudio para Argentina, Delfino (1988) especificó un modelo uniecuacional, lineal multivariable para analizar los resultados del aprendizaje en las escuelas primarias de Córdoba; considerando entre las variables explicativas un conjunto de factores asociados con la escuela (características del maestro y nivel de riqueza en el área de influencia) y con el hogar. El estudio encontró una asociación positiva y significativa aunque pequeña, entre los recursos docentes, la localización de la escuela y los resultados, pero no incluyó el control de las interacciones. Cervini (2002 a, 2002 b) estudió la equidad educativa en la Argentina y concluyó acerca de la existencia de desigualdades en el logro académico asociadas con las desigualdades socioculturales entre las familias de los alumnos, pero no incluye consideraciones acerca de las características que diferencian a las escuelas entre sí.

En los países de América Latina, de África y del sudeste asiático se desarrollaron en los años ochenta algunas aplicaciones del análisis de costos y efectividad para ayudar a evaluar programas de inversión en educación. El objetivo se centró, principalmente en identificar la contribución al aprendizaje de diferentes programas de inversión para mejorar las escuelas.

Estos programas adquirieron importancia en el contexto de la ayuda para el desarrollo y contaban con aportes de organismos multilaterales de crédito (capacitación docente, producción de textos escolares, radio-educación en zonas rurales, y similares). De Moura Castro y otros (1984) estimaron una función de producción para un conjunto de países de América Latina, a fin de analizar los determinantes del rendimiento escolar en la región. Lockheed y Hanushek, (1987) examinaron de forma comparada la relación de efectividad a costo para una variedad de programas de inversión en educación (radio-educación en Nicaragua, nuevos insumos para educación rural en Brasil, impacto del uso de textos escolares en México y otros, para una muestra de países en desarrollo. El estudio sirvió para examinar las dificultades que envuelve el uso del análisis comparado de situaciones educativas tan diversas como las contempladas. Mientras que Fuller (1986), en un análisis comparado de 30 programas de inversión observa que un mismo conjunto de inversiones (en entrenamiento de docentes, o en equipamiento de laboratorios, por ejemplo), luego de introducir controles por el origen social de los estudiantes, arroja resultados mixtos, aproximadamente en la mitad de los casos se indican efectos positivos mientras que en los restantes, los efectos resultan poco significativos o aún, de signo contrario. Hanushek (1996) aplicó técnica de meta-análisis a más de 100 estudios de efectividad-costos, arribando a una conclusión similar.

La variabilidad de los resultados en los estudios comparados es una consecuencia de las dificultades metodológicas que envuelve la recolección y tratamiento de cantidades importantes de información compleja sobre alumnos, docentes, padres, escuelas y otros aspectos. Pero las mayores dificultades para interpretar los resultados de los modelos lineales uniecuacionales son producidas esencialmente por la naturaleza jerárquica de los datos, al nivel de los alumnos, y en otros niveles, operacionalizar las variables del nivel del docente y de la escuela (que actúan sobre un grupo de alumnos) como agregados, pero que son incorporados luego como variables explicativas en modelos de un solo nivel para estimar la magnitud de sus efectos sobre la nota de cada alumno. Esto afecta el supuesto de independencia de las variables, y de independencia entre las variables y el error de estimación. Aún cuando el aprendizaje ocurre en el nivel del individuo, el hecho de que el aprendizaje ocurre en el aula y que las aulas forman parte de una escuela, que depende a su vez de unidades administrativas de orden superior demanda una especificación cuidadosa, que permita estimar la parte de la variabilidad en los resultados que puede ser explicada por la influencia de la pertenencia a determinado grupo. Las funciones de producción educativa uniecuacional no toman en cuenta el carácter esencialmente jerárquico de las relaciones que se pretenden estudiar y por esto han recibido fuertes críticas en los últimos años (Rowe, Hill y Holmes-Smith, 1995, Wenglinsky, 1998, Brik y Raudenbush, 1992, Monette et al, 2001).

Los modelos lineales jerárquicos permiten dar una respuesta preliminar a estas cuestiones porque están directamente orientados al tratamiento de datos anidados, como en el caso que nos ocupa, donde los estudiantes se agrupan por aula, las aulas pertenecen a escuelas y las escuelas responden a normativas y condiciones institucionales que varían entre jurisdicciones dando lugar a efectos sobre el rendimiento que se expresan, en definitiva, a diferentes niveles de agregación (Aitkin y Longford, 1986; Goldstein, 1984).

Cuando los datos corresponden a una estructura de este tipo, las observaciones individuales en la muestra no pueden ser tratadas como observaciones aleatorias y los métodos lineales sencillos, donde todas las variables explicativas entran en un mismo nivel producen estimadores sesgados e inconsistentes (Brik, Raudenbush, Seltzer y Congdon, 1986, Brik, y Raudenbush, 1992). La estructura lógica de los modelos jerárquicos permite agrupar a los

datos en dos, tres o más niveles, en función del problema a estudiar, la calidad de la información disponible y la capacidad de cálculo existente, mejorando la estimación de los efectos atribuibles a cada nivel de asociación de variables, permitiendo la descomposición de la variabilidad de la variable dependiente.

En este trabajo se desarrolla un modelo específico que identifica dos niveles de análisis. Un primer nivel, en el cuál se considera el conjunto de variables personales y del hogar de pertenencia que afectan el rendimiento de cada alumno. Un segundo nivel, contempla las variables cuyo efecto se expresa grupalmente sobre el conjunto de alumnos en cada aula (la calidad del maestro frente al curso, el equipamiento disponible, etc). El modelo plantea finalmente siete efectos fijos que capturan el efecto de la localización de las escuelas por regiones. Estos efectos fijos son interpretados como medidas del nivel relativo de la eficacia promedio de las escuelas en una región específica (el nivel de referencia restringe a cero el valor correspondiente a Ciudad de Buenos Aires). Las principales variables del modelo se detallan en la figura 1.

Figura 1
Las variables del modelo consideradas por nivel



El modelo explicativo de dos niveles que desarrolla este trabajo es el que describe la figura 1. Este modelo indica que el puntaje de cada alumno en una prueba estandarizada de cobertura nacional (en el área de Lengua, y de Matemática, separadamente) está

influenciado por variables que actúan al nivel del alumno (indicadas del lado izquierdo de la figura) y variables que corresponden al nivel del curso (indicadas del lado derecho). Las variables consideradas en el nivel del alumno se han clasificado en variables que caracterizan al alumno y variables que caracterizan su hogar. Entre las primeras está el sexo, la actitud hacia la asignatura, y la predisposición a aprender, evaluada por tres indicadores: si es repetente, la nota anterior en la asignatura, la nota promedio del curso anterior, excluida la asignatura. Las variables que caracterizan a su hogar incluyen (i) un índice de nivel socioeconómico (INSE) que combina educación del padre, condición de hacinamiento y posesión de bienes que reflejan el capital cultural y económico del hogar; (ii) tener hermanos desertores, (iii) posesión en el hogar de libros relacionados con la asignatura y (iv) nivel educativo de la madre.

Las variables del nivel curso, o escuela incluidas en el modelo permiten investigar separadamente los efectos atribuibles a las características del aula, a los atributos personales del docente, al tipo de gestión del establecimiento y a la localización de la escuela.

Las aulas se caracterizaron mediante (i) un indicador que expresa el capital cultural del grupo en relación al capital cultural promedio de todos los grupos-aulas- considerados; (ii) un índice de la calidad edilicia del aula (en base a datos sobre iluminación, ventilación y calefacción); nivel apreciado por el docente de la disciplina del curso; (iii) potencial uso de materiales pedagógicos (en base a la disponibilidad y estado de conservación); (iv) tamaño del curso (en base al número de alumnos que participaron de la prueba en cada curso).

Se incluyó separadamente una variable que permite diferenciar el efecto atribuible al tipo de gestión, considerando escuelas de gestión pública o privada.

También se investigó la influencia de los docentes evaluando cuatro características personales: (i) años de educación del docente, (ii) años de experiencia laboral, (iii) haber alcanzado la titularidad en el cargo, (iv) capacitación específica recibida en contenidos de su asignatura.

Mediante la inclusión de efectos fijos se evaluó el impacto de las diferencias regionales sobre el rendimiento escolar.

La figura 1 incluye, en su parte inferior, un conjunto de variables de interés, para las cuales no fue posible conseguir la información.

El modelo intenta dar una respuesta preliminar al siguiente conjunto de inquietudes:

¿Qué proporción de la variabilidad en las notas es directamente atribuible a las características personales del alumno y su hogar y cuánto de la misma es producto de diferencias entre escuelas en la calidad de la enseñanza?

¿Qué importancia relativa tiene el efecto de las condiciones del aula, los docentes y las escuelas para explicar la diferencia entre escuelas efectivas y las que no lo son tanto?

¿Qué diferencias se asocian con la región en la que residen?

¿Cuáles son los efectos más significativos y qué signo presentan?

¿Qué mecanismos y factores influyen sobre la efectividad del trabajo de los docentes en el aula?

En definitiva, el modelo propuesto permite indagar acerca de la efectividad de las escuelas, más específicamente, se intenta arribar a conclusiones preliminares que ilustren acerca de la eficiencia relativa con la que los diferentes factores y prácticas de gestión educativa promueven el mejor desempeño de los estudiantes en las pruebas escolares y si la región de pertenencia de la escuela afecta el nivel promedio de la nota.

3. LOS DATOS

Los datos provienen del Operativo Nacional de Evaluación correspondiente al año 2000 y están referidos a la totalidad de los alumnos de sexto grado que participaron a finales de ese año de las pruebas estandarizadas de matemática y de lengua que administra el Ministerio de Educación de la Nación. Conjuntamente con las pruebas, se suministra un formulario de auto-respuesta que recaba datos personales de los alumnos y su familia. Un segundo formulario es completado por el docente al frente del grado quien debe informar sobre una serie de aspectos que hacen a la calidad del aula, la disciplina, algunos rasgos personales que hacen a su condición laboral, y otras características. Un tercer formulario es suministrado al director del establecimiento y contiene preguntas de orden general sobre el rol del director, que no fueron incorporadas en este estudio. Para más detalles acerca de estas pruebas, las críticas que recibieron y el alcance de las mismas, véase Tenti Fanfani (2001).

a. Tipo de relevamiento y cobertura

El Operativo Nacional de Evaluación de la Calidad Educativa 2000 se desarrolló como relevamiento exhaustivo del conjunto de alumnos de establecimientos escolares que cursaban 6º grado del Primario o EGB. No obstante, hay algunos elementos que constituyen una excepción a la cobertura universal: en primer lugar, no se consideró a los establecimientos de la provincia de Neuquén y en segundo término, tampoco forman parte del relevamiento, y por lo tanto también son excluidos de los resultados obtenidos en este trabajo, los alumnos de establecimientos rurales con secciones de menos de 5 alumnos.

b. Información disponible

Las bases de datos del operativo han sido presentadas agrupando la información en tres niveles según su rol en el proceso educativo: alumnos, docentes y directivos. Estos niveles se corresponden con bases de datos independientes, con variables que permiten su vinculación (sección y establecimiento a las que concurre el alumno, sección por la que responde el docente y establecimiento al que pertenece el directivo).

La base de Alumnos contiene la información correspondiente al resultado de la evaluación de cada alumno en las asignaturas de Matemática y Lengua, así como datos provenientes de una encuesta que se administra de manera adicional (Cuadernillo Complementario del Alumno). El Cuadernillo contiene una indagación acerca de las características sociodemográficas del alumno, su hogar y la vivienda donde reside, la trayectoria escolar, opiniones sobre el ambiente familiar y escolar. A partir de este cuerpo de variables se trabajó para obtener la información del nivel alumno.

Tabla I: Distribución de los alumnos de 6to. grado evaluados por región

Región	Población (alumnos de establecimientos urbanos)				Muestra de trabajo utilizada (en %)			
	Alumnos		Rendimiento (en puntos*)		Alumnos		Rendimiento (en puntos*)	
	Cantidad	Porcentaje	Lengua	Matemática	Lengua	Matemática	Lengua	Matemática

Ciudad de Buenos Aires	36711	6.7	72.4	66.7	10.6	9.7	77.2	72.7
GBA	154542	28.3	61.2	56.1	24.3	23.0	70.7	65.9
Pampeana	180637	33.0	63.5	60.7	40.5	40.2	70.8	68.6
NEA	63389	11.6	57.1	53.5	8.5	8.5	62.9	60.2
NOA	49817	9.1	60.0	56.9	6.4	6.4	67.8	65.6
Cuyo	37728	6.9	63.9	61.3	7.6	8.4	70.9	69.2
Patagonia	23831	4.4	62.5	59.3	4.2	3.8	68.7	67.7
Total	546655	100	62.4	58.7	100	100	70.7	67.6
Total de casos en la muestra					n=55,165	n=64,482	n=55,165	n=64,482

Fuente: Elaboración propia con datos de ONE 2000 (Base de Alumnos). (*) Escala 0-100.

La base de datos de Alumnos reporta el resultado en las Evaluaciones de Matemáticas y Lengua para un total de 613,919 alumnos de 6º grado. De este total, 546.655 corresponden a alumnos de establecimientos urbanos, con la distribución por regiones detallada en la Tabla I. Tanto las bases de datos de alumnos como la de maestros adolecen de faltantes de información, esto obligó a realizar un detallado análisis de consistencia y selección de casos que determinó una muestra final de trabajo de 55165 casos en lengua (9% de la base total) y 64482 casos en matemática. (10.5% de la base total) con información completa para el conjunto de variables de interés (indicadas en la figura 1). La distribución de los valores de las variables en la muestra es representativa de los valores de las variables en la población, a nivel global y para cada una de las regiones que se analizan.

La población de varones y mujeres se divide en proporciones iguales; uno de cada cinco alumnos repite grado; tanto en lengua como en matemática los alumnos manifestaron una actitud positiva hacia el aprendizaje de la materia (en un índice de 0 a 5 la puntuación media se situó, en ambos casos, por encima de los 4 puntos con una desviación estándar de un punto); la nota promedio obtenida para la misma asignatura en el año anterior resultó llamativamente alta (8.3 en lengua y 8.2 en matemática) con una variabilidad de 2 puntos. En cuanto a los indicadores socioeconómicos, el hogar promedio pertenece a un estrato medio-bajo (el indicador señala una media de 40 puntos sobre un recorrido posible de cero a cien). Los maestros reportaron que las aulas presentan un estado satisfactorio de luz y ventilación, la disciplina del curso no es indicada como un problema serio para ellos, sí indicaron poseer escaso material didáctico y se obtuvo que alcanza a una media de 25 el número de alumnos (que rindieron la prueba) por curso. Los docentes, en promedio, indicaron poseer unos 15 años de educación y la desviación estándar es de apenas medio año. La experiencia promedio de los docentes es de 13 años, pero se observa una fuerte variabilidad (la desviación estándar es de 8 años). Un 75% de los docentes son titulares, pero solo 44% recibió capacitación específica en contenidos de su asignatura. El anexo con el detalle de los casos válidos por variable no se incluye por razones de espacio y puede ser solicitado a los autores.

4. EL MODELO DE ESTIMACION

La ecuación de estimación a dos niveles resultante surge de extender el procedimiento descrito en Raudenbush y Brik (1986) y en Goldstein (1986) que indican la conveniencia de plantear las ecuaciones lineales de cada nivel con sus respectivas variables, diferenciando los efectos fijos de los aleatorios. Las ecuaciones de cada nivel de análisis contienen los

términos de error atribuibles a cada uno de los efectos aleatorios considerados, y el residuo de la regresión estimada responde a los supuestos usuales de normalidad. Goldstein (1999) expresa el modelo multinivel utilizando directamente una única ecuación que contiene las variables de ambos niveles, junto con un término de error global que aparece descompuesto en ε_{ij} (error del nivel alumno, con distribución $N(0, \sigma_\varepsilon^2)$), y u_{kj} (error atribuido a las variables de nivel curso, con distribución $N(0, \sigma_{kk}^2)$). Ambos errores guardan independencia entre sí. Pero los u_{kj} pueden tener una estructura de correlación dada por una matriz $\tilde{\Sigma}$. En el análisis educacional, el supuesto corriente es que a la matriz Σ no se le impone una estructura predefinida (Singer, 1998).

$$Y_{ij} = \beta_{00} + \beta \cdot X_{ij} + \gamma \cdot Z_j + X_{ij} \cdot \lambda \cdot Z_j' + \mu' \cdot X_{ij} + \mu_{0j} + e_{ij} [1]$$

Donde:

Y_{ij} = variable dependiente medida en el primer nivel. Corresponde a la i-ésima observación ubicada en el j-ésimo grupo

β_{00} =intercepto

X_{ij} =vector de orden $k \times 1$ de variables medido en el primer nivel (i-ésima observación ubicada en el j-ésimo grupo). Se expresa como $(X_{1ij}, X_{2ij}, \dots, X_{kij})'$

β = vector de orden $1 \times k$ de parámetros expresado como $(\beta_{1j}, \beta_{2j}, \dots, \beta_{kj})$ y asociado con las variables del primer nivel

Z_j = vector de orden $p \times 1$ de variables medido en el segundo nivel (j-ésimo grupo). Presenta la forma $(1, Z_{1j}, Z_{2j}, \dots, Z_{p-1,j})'$

γ = vector de orden $1 \times p$ de parámetros (incluido el intercepto) asociado con las variables del segundo nivel

λ = matriz de orden $k \times p$ de parámetros simbolizada por λ_{ij}

y, donde

μ = vector de orden $k \times 1$ de efectos aleatorios para los coeficientes de las variables del primer nivel

μ_{0j} = desviación aleatoria del intercepto del j-ésimo grupo, respecto de la media general de Y_{ij} (o gran media) y

e_{ij} = término de error residual como es definido usualmente en el análisis de regresión.

Como surge de la ecuación [1], la misma contiene una parte de efectos fijos y otra de efectos aleatorios. Los parámetros que reflejan el efecto fijo y que deben ser estimados son: β_{00} (la ordenada al origen para el total de alumnos considerados en la estimación, con independencia del aula a la que pertenecen), los componentes de los vectores β (efecto fijo de las variables a nivel alumno), γ (efecto fijo a nivel curso) y λ (efecto fijo de las interacciones entre las variables del primer y segundo nivel). La parte aleatoria está compuesta por los errores arriba mencionados, cuyas varianzas y covarianzas, también son estimadas por el modelo.

La ecuación anterior fue corrida separadamente para los resultados obtenidos en las pruebas de Lengua y Matemática de Sexto grado, en el Operativo Nacional de Evaluación 2000 (ONE 2000), diferenciando las regiones del país de acuerdo a la clasificación del INDEC.

En el presente trabajo, el primer nivel de análisis corresponde al estudiante y el segundo al aula o curso. La variable explicada y_{ij} corresponde al puntaje obtenido por el alumno i en el aula j en la evaluación considerada (lengua o matemática). En cuanto a las variables explicativas empleadas, son las que corresponden al alumno y al aula, tal como está descrito en el diagrama de la figura 1. Todas las estimaciones fueron realizadas mediante la utilización del programa Stata 9.0.

5. RESULTADOS

En la Tabla 2 se presentan los resultados del análisis ANOVA realizado para identificar la importancia que adquiere la separación de los datos en dos niveles de análisis. La separación en niveles indicó que la contribución relativa del nivel 2, nivel curso, a la explicación de la variabilidad en los resultados es del orden de un tercio, observándose un valor de Rho algo superior en matemática relativo a Lengua (Rho=0.31 en Lengua y Rho=0.33 en Matemática). Esto manifiesta la presencia de un importante grado de anidamiento en los datos que torna desaconsejable utilizar MCO para las estimaciones.

Tabla 2

MODELO ANOVA

Rendimiento en la Asignatura	LENGUA		MATEMÁTICA	
	Estimación	Desv. Est.	Estimación	Desv. Est.
<i>Constante</i>	71.0792	490.26	67.8842	471.57
Varianzas				
Nivel alumno σ_{ϵ}^2	215.9219	1.50	235.3815	1.52
Intercepto σ_{00}^2	100.8708	2.39	118.4161	2.54
Número de observaciones	47572		55709	
Número de grupos	6363		7442	

Nota: El número total de observaciones (47572 y 55709) es inferior al que aparece en la Tabla 1 debido a que para el análisis anova y el posterior análisis multinivel se computaron las aulas de 4 o más alumnos
Fuente: Elaboración propia con datos de ONE 2000.

Las consecuencias de la presencia de anidamiento en los datos sobre las estimaciones pueden corregirse estimando una ecuación como la que se describe en [1] que recoge tanto efectos fijos como aleatorios en los coeficientes. La mencionada ecuación se aplicó a la estimación de los parámetros asociados con las variables que explican los rendimientos en las pruebas de lengua, por un lado, y en las pruebas de matemática, por el otro. Las covariables utilizadas como predictores se corresponden con las presentadas en el diagrama de la figura 1. En el anexo A se incluye una descripción detallada de las mismas, que

comprende: nombre, nivel de jerarquía, tipo, signo esperado del efecto, y construcción del indicador proxy. Los coeficientes fueron tratados como efectos fijos en el caso de las pendientes, mientras que se incorporó un efecto aleatorio a la ordenada al origen, o intercepto. Hay que destacar que en pruebas adicionales se incorporó también un efecto aleatorio en el coeficiente de la variable a nivel del alumno INSE (indicativa de nivel socioeconómico de su hogar), teniendo en cuenta la variable a nivel del curso SECTOR (indicativa del tipo de gestión estatal o privada). En la referida prueba, no resultaron significativas: la varianza de la pendiente del INSE, la covarianza entre ambas variables y la interacción, razón por la cual fue desestimada. La herramienta utilizada para correr este modelo fue el módulo XtMixed de Stata 9.0.

En la Tabla 3 se presentan los resultados de la estimación del modelo jerárquico en dos niveles para las pruebas de lengua y de matemática. Las variables se presentan agrupadas por nivel: alumno y curso. Dentro de las primeras, a su vez, se diferencian variables que aproximan el efecto individual y el efecto del hogar, mientras que, dentro de las segundas, se distinguen los factores asociados al aula, a las características del docente, a la escuela y a la región del país.

Tabla 3

MODELO SIN EFECTOS ALEATORIOS CON INTERACCIONES						
Rendimiento en la Asignatura (Lengua/Matemática)	LENGUA		MATEMÁTICA			
	Coef.	Est. Z	Coef.	Est. Z		
Nivel del Alumno						
<i>Propias del alumno</i>						
Sexo		Varón = 1	-4.4544	-32.71	1.8792	14.32
¿Repitió?		Si=1	-4.3980	-15.22	-3.7616	-13.34
Actitud hacia la materia	0.0123	4.88	0.0541	21.20		
Nota promedio en la materia en el año anterior	0.1065	22.93	0.1549	38.92		
Índice de notas en otras materias	0.1949	34.90	0.1793	35.09		
<i>Propias del hogar</i>						
INSE	0.0538	7.87	0.0706	10.54		
¿Deserción entre los hermanos?	-8.4451	-16.82	-7.3966	-15.04		
¿Hay libros de la asignatura en su hogar?	0.5369	2.86	(...)			
Educación de la madre	0.0067	2.50	0.0045	1.71		
Nivel del curso						
<i>Propias del aula</i>						
Estado del aula	0.0364	7.14	0.0463	8.84		
Disciplina del curso	0.0269	7.35	0.0206	5.32		
Disponibilidad y estado de materiales	0.0169	3.78	0.0170	3.31		
Número de alumnos por curso	0.3340	2.70	(...)			
Número de alumnos por curso al cuadrado	-0.0048	-2.04	(...)			
<i>Propias del docente</i>						
Educación del docente	0.3895	2.24	0.3714	1.91		
Experiencia del docente	0.1660	3.27	0.2176	4.10		
Experiencia del docente al cuadrado	-0.0025	-1.74	-0.0025	-1.70		
¿El docente es titular?	(...)		0.8026	2.40		
¿El docente tomó cursos de la asignatura?	0.4601	1.81	0.5816	2.23		
<i>Propias de la escuela</i>						
Sector		Privado = 1	5.8002	21.05	4.6756	16.05
<i>Propias de la región</i>						
		Ciudad de Buenos Aires = región base				
Cuyo	-3.9101	-7.22	-1.1123	-1.96		
Pampeana	-3.2304	-8.16	-1.0160	-2.36		
Nea	-7.2828	-11.59	-6.6724	-10.55		
Noa	-4.2385	-7.53	-2.6419	-4.54		
Patagonia	-3.6522	-5.63	(...)			
Partidos de GBA	-3.8639	-8.99	-3.6953	-7.87		
<i>Interacciones</i>						
INSE por Sector	(...)		(...)			
Sector por Repitencia	-1.6232	-2.31	-1.2343	-1.81		
Deserción por Repitencia	3.9068	4.30	4.0395	4.66		
<i>Constante</i>	60.4833	19.20	55.7409	16.34		
Varianzas						
Nivel alumno σ_{ϵ}^2	178.7315	1.25	196.8787	1.27		
Intercepto σ_{00}^2	51.5243	1.43	72.8475	1.73		
Número de observaciones	47572		55709			
Número de grupos	6363		7442			

Nota: La varianza de la pendiente se corresponde con la consignada arriba

La covarianza es entre el error de la pendiente y de la ordenada.

(...) Variables no significativas

*En negrita las variables significativas 10 por ciento

Fuente: Elaboración propia con datos de ONE 2000.

Con respecto a los resultados, en las variables propias del alumno se destaca el impacto que tienen el género y la repitencia. En cuanto al sexo, se encontró que el efecto de ser varón es negativo en la prueba de Lengua y de magnitud importante, del orden de -4.45 puntos. En matemática se observó el efecto contrario: ser varón incide en +1.88 puntos respecto de la mujer.

Por su parte, el efecto de la repitencia tiene una incidencia de -4.4 puntos en lengua y de -3.76 en matemática. Considerando la interacción con el tipo de gestión del establecimiento, el efecto llega a -6.02 y a -4.99, para lengua y matemática, respectivamente, cuando el alumno, además de ser repitente, asiste a escuelas de gestión privada. Este comportamiento aparentemente extraño, puede tener la siguiente explicación: los colegios privados normalmente poseen una proporción menor de alumnos repitentes, ya que tienden a expulsarlos (en la prueba se encontraron, aproximadamente, 13% de repitentes en establecimientos de gestión oficial y 2% en establecimientos de gestión privada). Sin embargo, cuando estas instituciones retienen a esos alumnos, su rendimiento está muy por debajo del promedio general. Un tema para investigar es, entonces, la posible segmentación del mercado entre establecimientos de gestión privada, según retienen o no a los alumnos repitentes

Cabe señalar que dentro de las variables propias del alumno también tienen efecto significativo sobre el rendimiento, aunque de menor impacto, tanto las notas obtenidas por el alumno en otras materias como en la misma materia, en años anteriores.

Con respecto a las variables propias del hogar, y siempre dentro del nivel del alumno, apareció como de fuerte impacto sobre los resultados la presencia de hermanos del alumno que abandonaron sus estudios, o nunca asistieron a la escuela. La regresión de Lengua muestra un coeficiente de -8.45 puntos asociados a esta variable, lo que debe ser interpretado como indicativo de un factor del entorno del hogar que opera negativamente sobre los resultados. En el caso de matemática se confirma este efecto, con un coeficiente de -7.4 puntos.

El índice de nivel socioeconómico, que varía en una escala 0-100, está asociado con un efecto unitario de 0.05, para el rendimiento en lengua, y de 0.07 para matemática. Un simple ejercicio numérico indicaría que, para la prueba de lengua, un alumno cuyo hogar se ubica un desvío estándar por debajo (o por encima) del INSE promedio (43.6), tiene un rendimiento en la prueba un punto por debajo (o por encima) del que obtendría si el INSE de su hogar fuera el promedio. Estos resultados contrastan con la opinión generalizada en el sentido de otorgar al factor socioeconómico del hogar un papel mucho más decisivo en la determinación de los resultados del alumno. Nótese que para hogares cuyo nivel socioeconómico oscila entre +/- un desvío estándar de la media, los resultados de las pruebas solo se modifican en no más de 2 puntos. Hay que agregar que al realizar la prueba del efecto de la interacción entre el INSE y la variable sector, que operacionaliza la gestión estatal o privada, tampoco aportó elementos que indiquen la existencia de una contribución significativa a la explicación de las notas obtenidas por el alumno. Todo esto llevó a concluir que hay espacio para un papel importante de la escuela en la promoción de mejores resultados.

El papel de la escuela fue analizado en el presente trabajo al nivel del curso, tal como se explicó más arriba, y al respecto, los principales resultados obtenidos son los que se comentan a continuación.

Los factores propios del funcionamiento del aula, tanto los edificios, como la disciplina del curso y los materiales de apoyo escolar mostraron una contribución interesante. Teniendo en cuenta que estas variables están expresadas en una escala de 0-100, el impacto diferencial sobre la nota promedio de lengua de un alumno que asiste a un aula que, por su estado, se encuentra ubicada en la posición 75 de la escala, es de +2 puntos en relación a otro alumno con similares características, pero que asiste a un aula ubicada en la posición 25 de dicha escala. En matemática, el impacto del estado del aula es levemente superior, ya que para la misma comparación la diferencia asciende a +2.5 puntos. Similares consideraciones cabe realizar con relación a la disciplina del curso y la disponibilidad y estado de materiales didácticos. Con los datos disponibles, no fue posible, en cambio, obtener resultados que permitan conocer la importancia del tamaño del curso en la determinación de las notas de los alumnos. Existe la presunción que la cantidad de alumnos que se presentaron a rendir la prueba no constituye un indicador apropiado del verdadero tamaño del aula, debido a que no existió posibilidad de controlar las inasistencias.

La contribución del docente frente al aula se trató de captar mediante 4 factores. El efecto asociado con la variable nivel de escolaridad alcanzado por el maestro, evaluada en el promedio, es del orden de +5.8 puntos, en lengua, y de +5.5 puntos en matemática. Por su parte, la experiencia del docente, computada en el promedio, impacta en +2.2 puntos de las pruebas de lengua y 2.9 de las pruebas de matemática. Un año más de educación del docente posee un impacto reducido, de menos de medio punto. Mientras que, diez años más de experiencia en relación al promedio implican +2.2 puntos en los resultados de lengua y +2.9 en los resultados de matemática. Si el docente tomó algún curso relacionado con los contenidos de su asignatura, se reflejó en un efecto positivo sobre la nota de sus alumnos que alcanzó a +0.46 para lengua y +0.58 para matemática. Nótese que el impacto de esta última variable resulta superior al asociado con aumentar en un año la escolaridad del docente (0.39 en lengua y 0.37 en matemática). Con lo cual habría que prestarle más atención a la relación costo-efectividad asociada con extender los programas de capacitación específica vs extender la educación general.

El efecto de la gestión privada de los establecimientos educativos es positivo sobre los resultados de las pruebas de los alumnos, ya que la regresión mostró que en lengua el puntaje en estas instituciones supera en 5.8 al de las escuelas de gestión estatal. En matemática, el efecto sobre los resultados es menor, ubicándose en +4.7 puntos.

Por último, se introdujo una variable dummy para tratar de captar el efecto diferencial sobre las notas atribuibles a la región del país donde se encuentra ubicado el establecimiento escolar. Para ello se utilizó la regionalización de INDEC. Un primer aspecto a destacar está relacionado con la diferencia encontrada en el orden de magnitud de los coeficientes de la regresión asociados con cada región, pudiéndose observar que el efecto es mayor en lengua que en matemática. Otro aspecto, es la variabilidad del nivel del efecto entre regiones. Por ejemplo, tomando como referencia un alumno promedio, si estudia en el NEA, se espera que obtenga, en la prueba de lengua, 7.28 puntos menos de los que habría logrado si residiera en la ciudad de Buenos Aires. Para la prueba de matemática, el puntaje sería inferior en 6.67. Esta es la región que representa el caso extremo, seguido por el NOA. La región pampeana, en cambio, es la que se encuentra más próxima en resultados de ambas pruebas, a la ciudad de Buenos Aires (-3.2 en lengua y -1.0 en matemática).

Con el objeto de mostrar las mejoras en las estimaciones de los efectos resultantes de aplicar la regresión multinivel, con relación a los resultados que se obtienen de aplicar MCO, incorporamos en la Tabla 4 las estimaciones que surgen al aplicar este último método.

Tabla 4

MODELO MCO

Rendimiento en la Asignatura (Lengua/Matemática)	LENGUA		MATEMÁTICA		
	Coef.	Est. Z	Coef.	Est. Z	
Nivel del Alumno					
<i>Propias del alumno</i>					
Sexo	Varón = 1	-4.5892	-32.12	1.896	13.46
¿Repitió?	Si=1	-5.4213	-17.85	-5.139	-16.99
Actitud hacia la materia		0.0084	3.19	0.050	18.12
Nota promedio en la materia en el año anterior		0.0985	19.88	0.146	33.68
Índice de notas en otras materias		0.1974	33.20	0.180	32.45
<i>Propias del hogar</i>					
INSE		0.1037	14.90	0.131	18.87
¿Deserción entre los hermanos?		-9.8619	-18.32	-8.823	-16.34
¿Hay libros de la asignatura en su hogar?		0.8435	4.26	(...)	
Educación de la madre		0.0133	4.64	0.013	4.55
Nivel del curso					
<i>Propias del aula</i>					
Estado del aula		0.0336	10.39	0.039	12.39
Disciplina del curso		0.0254	11.06	0.020	8.85
Disponibilidad y estado de materiales		0.0150	5.39	0.014	4.52
Número de alumnos por curso		0.3112	3.96	(...)	
Número de alumnos por curso al cuadrado		-0.0045	-3.13	(...)	
<i>Propias del docente</i>					
Educación del docente		0.4041	3.77	0.301	2.65
Experiencia del docente		0.1517	4.81	0.191	6.09
Experiencia del docente al cuadrado		-0.0025	-2.79	-0.002	-2.26
¿El docente es titular?		(...)		0.837	4.19
¿El docente tomó cursos de la asignatura?		(...)		0.525	3.40
<i>Propias de la escuela</i>					
Sector	Privado = 1	4.7631	26.33	3.454	19.06
<i>Propias de la region</i>					
Cuyo	Ciudad de Buenos Aires = región base	-3.4025	-10.17	(...)	
Pampeana		-2.8068	-11.53	-0.216	-0.87
Nea		-6.7474	-16.96	-5.783	-15.34
Noa		-3.6913	-10.52	-1.602	-4.65
Patagonia		-3.0962	-7.49	0.935	2.17
Partidos de GBA		-3.1449	-11.77	-2.698	-9.83
Interacciones					
INSE por Sector		(...)		(...)	
Sector por Repitencia		-1.4489	-1.94	(...)	
Deserción por Repitencia		4.8562	4.99	4.636	4.86
<i>Constante</i>		60.5815	30.61	56.542	27.87
R²		0.2772		0.2441	
Número de observaciones		47572		55709	

(...) Variables no significativas

*En negrita las variables significativas 10 por ciento

Fuente: Elaboración propia con datos de ONE 2000.

Comparando las tablas 3 y 4, se aprecia que hay diferencias en los valores de los coeficientes así como en sus errores estándar, relacionados con la eliminación de sesgos y el aumento de la eficiencia de la estimación producto de neutralizar el efecto cluster. A este respecto es interesante ilustrar las diferencias en una de las variables centrales en esta

discusión, tal como es el efecto asociado con nivel socioeconómico del hogar, reflejado a través del INSE. Los valores de los coeficientes obtenidos por MCO fueron 0.1037 en lenguas y 0.131 en matemáticas, mientras que en las estimaciones por multinivel, los coeficientes resultaron 0.054 y 0.071, respectivamente. La disminución del impacto se reduce prácticamente a la mitad, en concordancia con resultados obtenidos en trabajos similares para otros países que indican que la utilización de MCO sobreestima la importancia del hogar en detrimento del rol de la escuela. El efecto de tener hermanos desertores también se reduce al aplicar multinivel, pero el cambio es menor.

En contraposición, el efecto asociado con las variables de la escuela se refuerza en el análisis multinivel, en comparación con la regresión de MCO. A título de ejemplos, la incidencia de la gestión pasa de un coeficiente de 4.8 a 5.8 en lenguas, y de 3.5 a 4.7 en matemáticas, y el efecto de la experiencia del docente por año adicional, aumenta de 0.15 a 0.19 en lenguas, y de 0.16 a 0.22 en matemáticas.

Por su parte, en la Tabla 5, se presenta la descomposición de la variabilidad en ambas regresiones, con la finalidad de mostrar la ganancia en la explicación asociada con el uso de multinivel.

Tabla 5

Lengua

Modelo	R² alumno	R² curso	R²
MCO			0.277
Modelo Multinivel	0.172	0.489	0.273

Matemática

Modelo	R2 alumno	R2 curso	R2
MCO			0.244
Modelo Multinivel	0.164	0.385	0.238

Fuente: Elaboración propia con datos de ONE 2000.

Las variables del nivel alumno contribuyen a explicar el 17.2% de la variabilidad de ese nivel en lengua, y 16.4% en matemática. Mientras que el aporte de las covariables asociadas específicamente al nivel curso aportan un 48.9% a la explicación en el caso de lengua y 38.5% en matemáticas.

6. CONCLUSIONES

El análisis jerárquico posee ventajas en el estudio del rendimiento escolar de los alumnos, en relación con la utilización de técnicas tradicionales de estimación uninivel, en la medida que permite neutralizar parcialmente el efecto anidamiento de los datos.

Una aplicación de esta técnica al caso de Argentina con datos del Operativo Nacional de Evaluación de la Calidad Educativa del año 2000, permitió corregir parcialmente la sobreestimación del efecto atribuible a las características del hogar y del alumno sobre los resultados de las pruebas de lengua y matemática, que surge cuando se aplica estimación uninivel. A tal fin se utilizó una modelización que incluyó covariables con efectos fijos y un efecto aleatorio sobre el intercepto.

Particularmente, se observó una disminución importante en el papel del nivel socioeconómico del hogar y de las variables asociadas con el alumno, y un aumento correlativo en la importancia de las variables del curso. Tanto el tipo de gestión de la escuela, como las condiciones de calidad del aula y el título y experiencia del docente resultaron, todos ellos, factores de relevancia para explicar el rendimiento. La región del país donde se encuentra ubicado el curso fue también un factor de impacto importante, reflejando las diferencias conocidas en las condiciones socioeconómicas generales en el territorio nacional. Estos aspectos deberían ser cuidadosamente observados al momento de definir políticas públicas para el sector.

Bibliografía

- Aitkin M, Longford, N. (1986). Statistical Modelling Issues in School Effectiveness Studies (with discussion). *Journal of the Royal Statistical Society. Series B* V. **149**, Pp. 1-43.
- 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 (2002a). Desigualdades en el logro académico y reproducción cultural en Argentina. Un modelo de tres niveles. *Revista Mexicana de Investigación Educativa* V. **7**, Pp. 445-500.
- 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.
- Coombs PH, Hallak, J. (1987). *Cost Analysis in Education: A Tool for Policy and Planning*. The John Hopkins University Press, Baltimore y Londres.
- De Moura Castro C, Sanguinety, J. A., Marques, E. A., Rodriguez Lacerda, E., Ciavatta Franco, M. A., Silva, M. A. (1984). *Determinantes de la Educación en America Latina: Acceso, Desempeño y Equidad*. Editora da Fundação Getulio Vargas, Rio de Janeiro.
- Delfino JA (1989). Los Determinantes del Aprendizaje. *Ensayos en Economía de la Educación*. ed. Editor Hp, pp. 287-316, Buenos Aires.
- Fuller B (1986). Raising School Quality in Developing Countries: What Investments Boost Learning? *World Bank Discussion Papers*, pp. 1-81. The World Bank, Washington D.C.
- Goldstein H (1986). Multilevel Mixed Linear Model Analysis using Iterative Generalized Least Squares. *Biometrika* V. **73**, Pp. 43-76.
- Goldstein H (1991). Multilevel Modelling of Survey Data. *The Statistician* V. **40**, Pp. 235-244.
- Goldstein H (1999). *Multilevel Statistical Models*. ed. London Ioeo.
- Goldstein H, Yang, Min, Omar, Rumana, Turner, Rebeca, Thomson, Simon (2000). Meta-Analysis Using Multilevel Models with an Application to the Study of Class Size Effects. *Applied Statistics* V. **49**, Pp. 399-412.
- Hanushek E (1994). *Making Schools Work: Improving Performance and Controlling Cost*. The Brookings Institution, Nueva York.

- Hanushek E (1996). School Resources and Student Performance. In *Does Money Matter? The Effect of School Resources on Student Achievement and Adult Success*. ed. Burtless G. Brooking Institution Press, Washington.
- 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.
- Herrero V., Palacios P., Ruíz Díaz, F. (2005) Un índice de nivel socio-económico de los hogares para aplicar a un modelo explicativo del rendimiento escolar. *XXXIII Coloquio de la Sociedad Argentina de Estadística*. Villa Giardino. Córdoba. Argentina.
- Hox JJ Multilevel Modeling: When and Why. *Classification, Data Analysis, and Data Highways*. ed. I. Balderjahn Rmams, Pp. 147-154. Springer Verlag, New York.
- Hox JJ (1995). Applied Multilevel Analysis, pp. 126. TT-Publikaties.
- Johnes G (1995). Eficiencia Interna y Calidad de la Educación. *Educación para el Desarrollo*, pp. 195-227. Ministerio de Trabajo y Seguridad social.
- Leigh A, Simmons, John (1975). The Determinant of School Achievement in Developing Countries: The Educational Production Function. *International Bank for Reconstruction and Development*, pp. 1-66.
- Levin H (1970). A Cost-Effectiveness Analysis of Teacher Selection. *Journal of Human Resources* **V. 5**, Pp. 34-33.
- Levin H (1983). *Cost-Effectiveness*. Sage Publications, Nueva York.
- Levin H, Bowels, S. (1968). The Determinants of Scholastic Achievement. An Appraisal of some recent Evidence. *Journal of Human Resources* **V. 3**, Pp. 3-24.
- Levin H, Glass, J., Meister, G. (1984). *Cost-Effectiveness of four Educational Interventions*. Stanford School of Education.
- Levin HM (1971). A New Model of School Effectiveness, pp. 55-75. Secretary of HEW, Washington D.C.
- Lockheed Marlaine E., Hanushek Eric (1987). Improving the Efficiency of Education in Developing Countries: Review of the Evidence. The World Bank, Report No. EDT77.
- Monette G, Shao, Qing and Kwan, Ernest (2001). A First Look at Multilevel Models, pp. 96. Institute for Social Research.

- 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 N° 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.
- Tenti Fanfani, Emilio, Comp. (2001) El rendimiento escolar en la Argentina, Buenos Aires Editorial Losada.
- Wenglinsky H (1998). *School District Expenditures, School Resources and Student Achievement: Modeling the Production Function*. Education Testing Service, Washington.

Anexo A

Tabla de definiciones y construcción de variables

VARIABLES	SIGNO ESPERADO	TIPO	CONSTRUCCIÓN (a partir del cuestionario del alumno y el docente de las bases ONE 2000)	REGRESIÓN
Nivel del alumno				
<i>Propias del alumno</i>				
Sexo	(+) en matemática y (-) en lengua	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.	
Actitud hacia la materia	(+)	0-100 (5 niveles)	Se calculó un índice a partir de qué piensa el alumno acerca de matemática (preg. 17) y lengua (preg. 18), asignando un 1 si estaba de acuerdo con afirmaciones positivas hacia la materia y 0 si estaba en desacuerdo. El índice se creó mediante el cociente entre la suma de las respuestas dadas por el alumno a 4 afirmaciones y la cantidad de respuestas sumadas en el numerador. El rango fue llevado de 0 a 100.	Variable centrada respecto a su media (*)
Nota promedio en la materia en el año anterior	(+)	0-100 (4 niveles)	Consiste en la nota en el período t-1 del área particular que se está analizando, recodificando la preg. 21 con 7.5 (notas 10, 9 y 8), 5 (7 y 6), 2.5 (5 y 4) y 0 (3, 2 y 1) y transformando luego el rango de 0 a 100	Variable centrada respecto a su media
Índice de notas en otras materias	(+)	0-100 (10 niveles)	Se creó el índice mediante el cociente entre la suma de las notas de tres materias (Lengua, Matemática, Ciencias Sociales y Naturales, exceptuando la correspondiente a la regresión que se desea realizar) y la cantidad de notas sumadas en el numerador. Se llevó el rango de 0 a 100.	Variable centrada respecto a su media
<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

VARIABLES	SIGNO ESPERADO	TIPO	CONSTRUCCIÓN (a partir del cuestionario del alumno y el docente de las bases ONE 2000)	REGRESIÓN
Disciplina del curso	(+)	0-100 (6 niveles)	Se creó a través de la pregunta "¿Cuántos de sus alumnos son muy disciplinados?" (p33_5) del cuestionario del docente, mediante el cociente entre la suma de las respuestas dadas por el docente y la cantidad de respuestas sumadas en el numerador. El rango se llevó de 0 a 100.	Variable centrada respecto a su media
Disponibilidad y estado de materiales	(+)	0-100 (continua)	Se creó a través de la preg. 41 del cuestionario del docente, la cual indaga acerca de la disponibilidad de materiales en la escuela, el estado de los mismos y con qué frecuencia son utilizados por el docente. Se creó el índice mediante el cociente entre la suma de las respuestas dadas por el docente en forma de condición (1=dispone y 0=no dispone; si dispone, 1=estado bueno, 0,5=regular y 0=estado malo; y para cada una de estas desagregaciones, 1=uso frecuente, 0,5=algunas veces, 0=casi nunca) y la cantidad de respuestas sumadas en el numerador.	Variable centrada respecto a su media
Número de alumnos por curso	(-)	(1-60)	Indica la cantidad de alumnos que rindieron las pruebas del ONE de una misma sección de un determinado establecimiento.	Se impuso como condición que tuviera en cuenta sólo los cursos con 10 alumnos o más y 45 alumnos o menos.
Número de alumnos por curso al cuadrado	(-)	(1-3600)	Se elevó la variable anterior al cuadrado con el fin de ver si había una cantidad óptima de alumnos por curso.	La condición anterior restringe el rango de 100 a 2.025
<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 construidos en años que van de "menos de un año" a "más de 30 años" con saltos cada 5 años.	
Experiencia del docente al cuadrado	(-)	0.25-1225 (8 niveles)	Se elevó la variable anterior al cuadrado con el fin de ver si con una cierta cantidad de años de experiencia la enseñanza comenzaba a empeorar.	
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.	
<i>Propias de la región</i>				
Cuyo	(-)	Dummy	Se construyeron con referencia a Capital Federal, con el fin de ver si existía una diferencia significativa en las notas obtenidas por los alumnos pertenecientes a las distintas regiones en relación a la que obtiene un estudiante que pertenece a la Capital del país.	
Pampeana	(-)	Dummy		
NEA	(-)	Dummy		
NOA	(-)	Dummy		
Patagonia	(-)	Dummy		
Partidos del GBA	(-)	Dummy		
Interacciones				
Inse por sector	?		Producto de variables indicadas	
Sector por repitencia	?		Producto de variables indicadas	
Deserción por repitencia	El efecto se potencia		Producto de variables indicadas	

Notas (*) Todas las variables continuas fueron centradas en torno a la media general de manera de facilitar la interpretación de la ordenada al origen, que resulta ser el promedio correspondiente a los alumnos que cumplen con un determinado perfil.

(**) La metodología para la construcción del índice puede ser consultada en Herrero, Palacios y Ruiz D. (2005).