

FACTORES ASOCIADOS AL LOGRO ESCOLAR / 5

**Calidad y equidad en la
Educación Básica de
Argentina**

**Alumnos de 7° año - Escuela urbana
Rendimiento en Matemática
Operativo Nacional de Evaluación 1997**

Ministerio de Cultura y Educación de la Nación
Secretaría de Programación y Evaluación Educativa

Presidente de la Nación
Dr. Carlos Saúl Menem

Ministro de Cultura y Educación
Dr. Manuel García Solá

Secretario de Programación y Evaluación Educativa
Prof. Sergio Luis España

Subsecretario de Evaluación Educativa
Lic. Pablo Narvaja

Subsecretaria de Gestión Educativa
Lic. Irene Kit

Director del Instituto Nacional de Educación Técnica
Prof. Carlos Palacio

Directora Nacional de Evaluación
Lic. María Lucrecia Tulic

**Calidad y equidad en la
Educación Básica de
Argentina**

por

Rubén Cervini

Ministerio de Cultura y Educación de la Nación
Buenos Aires, Noviembre 1999

PRESENTACIÓN

El Ministerio de Cultura y Educación de la Nación, a través del Sistema Nacional de Evaluación de la Calidad, ha implementado anualmente, desde 1993, el Operativo Nacional de Evaluación (ONE), cuyo **objetivo principal** es diagnosticar el nivel de aprendizaje de los alumnos del nivel básico. Durante esas evaluaciones, se han aplicado pruebas de opción múltiple a muestras de alumnos de 3º, 6º y 7º años de la EGB en diferentes área curriculares.

Junto a este objetivo central, en el SINEC ha estado presente la idea de que es necesario conocer qué variables inciden sobre el nivel y la distribución de esos aprendizajes. Conocer esos factores es de máxima relevancia para el sistema educativo ya que, sobre esa base, se podrán formular y evaluar políticas y programas educativos y al mismo tiempo, promover el compromiso y la participación de los diversos actores del sistema educativo.

Como una forma de concretar esta idea, en todas las evaluaciones realizadas desde 1993, se elaboraron y aplicaron cuestionarios para los alumnos, los maestros y los directores, a través de los cuales se relevaron informaciones orientadas a mejorar la comprensión del comportamiento del rendimiento de los alumnos en las pruebas, es decir, a identificar algunos de los factores del logro escolar.

El **objetivo de esta publicación** es divulgar los resultados obtenidos al analizar algunas informaciones provenientes de los cuestionarios aplicados al alumno de 7º año, al director y a los docentes del establecimiento, durante el ONE de 1997, referidas principalmente, a los llamados "factores extra-escolares". Más específicamente, el trabajo se propone profundizar el conocimiento del comportamiento del rendimiento del alumno en relación con su origen social y algunas de sus características personales, por un lado, y con la composición social, la dotación de recursos y el sector de dependencia (público/privado) de la escuela a la que asiste, por el otro; es decir, se apunta a develar la incidencia que tienen algunos aspectos del Contexto educativo (factores extra-escolares) y la disponibilidad de Insumos escolares (factor escolar), sobre el nivel y distribución del aprendizaje de los alumnos. Para alcanzar este objetivo, el análisis se desarrolla en base a tres conceptos claves, a saber: **calidad, equidad y eficacia educativa.**

En otra publicación de esta misma serie, se divulgan informaciones referidas a las relaciones entre el rendimiento y diversas características institucionales de la escuela y de sus prácticas pedagógicas. La identificación de estos factores es el resultado de diversos análisis estadísticos realizados por el equipo técnico responsable desde el inicio del SINEC en 1993.

El presente trabajo se inicia con la exposición de algunos elementos metodológicos necesarios para comprender su contenido (Capítulo 1). Inmediatamente después (Capítulo 2), se analiza la (des)composición de la variación del rendimiento en sus dos componentes: la variación "entre-escuela" y la variación "entre-alumnos", punto de partida para evaluar, a continuación, el grado de equidad y calidad educativa detectada a través de los datos (Capítulo 3). Sobre esa base, en el Capítulo 4 se determina la asociación entre el rendimiento y algunas características personales

objetivas del alumno (edad, sexo, repitencia, cambio de escuela), mientras que en el Capítulo siguiente, el foco de interés son las variables actitudinales del alumno.

Se pasa luego a estudiar el efecto de los recursos escolares disponibles en el establecimiento (Capítulo 6), para finalmente, abordar el análisis de la diferencia de aprendizaje entre los sectores público y privado (Capítulo 7), tomando en cuenta las variables contextuales y de insumo, es decir, con una perspectiva de equidad educativa.

Este trabajo fue realizado
con el apoyo del siguiente
equipo técnico:

Dra. Basualdo, Marisa
Lic. Flores, Susana
Prof. Izaguirre, Luisa

Procesamiento

Ing. Baruzzi, Graciela
Lic. Riera, Estela

Las opiniones emitidas por el autor de este trabajo son
de su exclusiva responsabilidad y no comprometen
al Ministerio de Cultura y Educación de la Nación.

INDICE

	Página
Capítulo 1. Metodología	1
Capítulo 2 . La distribución y las desigualdades del rendimiento en Lengua y Matemática.	5
Capítulo 3. La equidad educativa	13
Capítulo 4. Rendimiento y características personales del alumno	31
Capítulo 5. Actitudes, representaciones y rendimiento del alumno	45
Capítulo 6. Los recursos escolares	57
Capítulo 7. Escuela pública y privada.	69
Anexo . Cuestionarios del Alumno, del Director y del Docente	79

Capítulo 1

METODOLOGIA

El enfoque metodológico adoptado para lograr el objetivo propuesto inicialmente se sitúa dentro de la tradición de estudios estadísticos "correlacionales". La determinación del comportamiento del rendimiento y la identificación de sus factores se realiza a través del análisis del grado de asociación entre variables "independientes" o "explicativas" (los factores potenciales) y la variable "dependiente" o de respuesta (rendimiento en Matemática).

Una advertencia inicial. Siempre parece justificado recordar una advertencia epistemológica muy común en este tipo de estudios: encontrar una asociación estadística no significa necesariamente haber explicado el fenómeno, en nuestro caso, la variación del rendimiento; para hallar un por qué no sólo es necesario una observación detallada de los "micro-procesos" subyacentes, sino también una teoría que la ilumine.

Por ello, cuando a lo largo de este capítulo se hable de "efecto", "explicación", "determinación", etc., en realidad nos estaremos refiriendo y restringiendo a la asociación (correlación) estadística entre variables, o más particularmente, a **la capacidad predictiva de alguna/s variable/s respecto al rendimiento en Matemática.**

Fuentes de información. Los datos analizados en este trabajo provienen de la prueba de Matemática y de los cuestionarios del alumno (AP7), del Director (DIP) y del docente (DOP). Mientras que la casi totalidad de las informaciones contenidas en el AP son incorporadas en el análisis, del DIP y el DOP sólo se consideran las informaciones referidas a la dotación de recursos en la escuela. En Anexo se presenta un ejemplar del AP y las preguntas del DIP y del DOP que son utilizadas.

Aplicación de los instrumentos. Todos los cuestionarios son auto-aplicados. En general, los alumnos responden el cuestionario durante el primer día del operativo de evaluación y después de la aplicación de la prueba de Matemática. Por razones operacionales, sin embargo, en algunas jurisdicciones y/o escuelas el cuestionario puede haberse aplicado el segundo día. El director responde al DIP en cualquier momento durante los días que dura el operativo de evaluación. El maestro en cambio, lo responde en el momento en que se está aplicando la prueba a sus alumnos.

Definición y tipos Variables. La definición operacional de cada variable se presenta en el capítulo donde se la analiza. Las variables incluidas en este estudio pueden ser clasificadas según nivel de agregación, forma de construcción y contenido:

- a) agregación:
 - . alumno; ej: rendimiento;
 - . escuela; ej: dotación de recursos;

- b) construcción:
 - . simples; ej. género del alumno;
 - . compuestas; ej. nivel socioeconómico;

- c) contenido:
 - . objetivas; ej: edad del alumno;
 - . subjetivas; ej.: motivación del alumno.

La muestra. Para la aplicación de las pruebas, el ONE implementa anualmente un diseño muestral de tipo estratificado y por conglomerados. Los estratos son la localización (urbano, rural), la jurisdicción y el régimen (público/privado). Los conglomerados son las secciones de alumnos. La prueba se aplica a todos los alumnos que integran cada sección seleccionada. La muestra de los cuestionarios para el alumno está condicionada por la muestra de las pruebas ya que se aplica a todos los alumnos que son evaluados.

Del total de la muestra de la ECE/97, en este trabajo se analizan solamente las secciones de 7° año del área urbana que cuenten con 20 o más alumnos con informaciones válidas en las variables consideradas. Esta condición responde a exigencias de la técnica de análisis estadístico utilizada (ver abajo), para conseguir estimaciones confiables. Debido a la variación de la cantidad de "missing" (no respuestas) por variable, el tamaño muestral varía de acuerdo a las variables que se estén analizando; por lo tanto, se lo indica en cada uno de los modelos analizados.

La técnica de análisis. Para el análisis de las relaciones entre el rendimiento y las diferentes variables consideradas, se aplicó la técnica de "análisis estadístico de niveles múltiples"¹. Esta es una técnica correlacional adecuada para analizar variaciones en las características de los individuos que son miembros de un grupo, o sea, mediciones que forman parte de una estructura agrupada y jerárquica².

Este es el caso más frecuente para las mediciones en educación. Los alumnos son parte de un grupo ("el aula", "la sección A de séptimo año"), que pertenece a una "escuela", la cual se encuentra en un "distrito" de una "jurisdicción", etc. Este es precisamente, el tipo de datos analizados en este trabajo.

Los alumnos pertenecientes a una misma escuela participan homogéneamente de algunas características (por ejemplo, contexto socioeconómico) y simultáneamente, todos ellos se diferencian por igual de los alumnos de otra escuela (con un contexto socioeconómico eventualmente diferente). En ese tipo de realidad, para explicar la variación de los comportamientos individuales (el rendimiento escolar del alumno, por ejemplo), se deberán investigar no sólo las características del propio alumno (ejemplo: nivel socioeconómico familiar del alumno), sino también las del grupo escolar del que éste forma parte (ejemplo: composición social de la escuela). En otras palabras, **los factores del aprendizaje deberán ser especificados por nivel de agregación.**

¹ Para un mayor detalle de esta técnica, ver Anexo B de "Los Factores del rendimiento en la Educación Primaria –1995-", Rubén Cervini y otros, en **Informe de Investigación N° 25**. Ministerio de Cultura y Educación de la Nación.

² El término jerárquica es usado aquí para denotar la inclusión de las observaciones en niveles de creciente agregación. El término "agrupamiento" se refiere al hecho de que los alumnos son parte del sistema educativo a través de su pertenencia a un grupo (división, sección, escuela), y no de forma aislada. En el caso de la ECE/97, sección y escuela son equivalentes.

Al igual que en un modelo de regresión tradicional, será imposible explicitar, en cada nivel de agregación (alumno, sección, escuela, etc.), todas las influencias sobre el comportamiento individual. Así por ejemplo, habrá una porción de la variación del rendimiento del alumno que se debe a las características del grupo al que pertenece pero que no se pueden incluir explícitamente. Por lo tanto, esa variación no-explicada debería ser considerada como parte del error aleatorio en ese nivel de agrupamiento. El análisis de niveles múltiples presta atención justamente a esa situación, permitiendo incorporar determinadas características del grupo dentro del modelo y produciendo estimaciones correctas de errores estándar. Dicho de otra forma, la técnica permite que las relaciones individuales varíen entre los grupos, evaluando entonces, el efecto del agrupamiento sobre los errores estándar. Sobre esta base, será posible definir los tests estadísticos y los intervalos de confianza correspondientes para evaluar los efectos de cada variable.

En la especificación de cualquier modelo con esta técnica, es posible definir 2 partes:

. **Parte Fija:** son los parámetros que permiten definir una línea promedio para **todos** los alumnos de **todas** las escuelas.

* Por ejemplo, si estamos estudiando la relación entre el rendimiento y el nivel socioeconómico del alumno, en la **parte fija** del modelo se define:

. el valor promedio (ajustado) del rendimiento, es decir, el valor donde la línea que representa la relación entre ambas variables corta a la ordenada (rendimiento), y

. el valor promedio de la fuerza de la relación entre rendimiento y nivel socioeconómico, es decir, la pendiente de la línea que representa la relación entre ambas variables.

. **Parte Aleatoria:** muestra, en cada nivel de agregación, la estimación de la variación de los parámetros determinados en la parte fija.

* En nuestro ejemplo de arriba, en la parte aleatoria del nivel 2 (escuela) aparecen las estimaciones de la variación de las líneas individuales de las escuelas en torno de la línea promedio general y de la variación de los rendimientos observados de los alumnos alrededor de la línea promedio de su escuela.

En resumen, los principales atractivos de esta técnica radican en que ofrece la posibilidad de:

(a) modelar simultáneamente los diferentes niveles de variación (por ejemplo, alumno y escuela), permitiendo, por tanto, saber qué proporción de la variación del rendimiento escolar se debe a características del alumno (nivel 1) y cuál a características de la sección o escuela³ (nivel 2), y

(b) permitir que el nivel de rendimiento (intercepto a) y la fuerza de relación o interacción entre los factores (pendiente b) varíe libremente en los diferentes niveles de agregación (alumno, secciones o escuelas).

³ En términos técnicos, se trata de identificar la varianza del rendimiento entre alumnos (“dentro de la escuela”) y la varianza entre las diferentes secciones (“entre escuela”).

Dado que esta técnica de análisis no es de conocimiento común en nuestro medio, cada vez que, a lo largo de este trabajo, se realice una nueva operación, ésta será precedida por una breve explicación de su significado.

Niveles de análisis. Los datos de la ECE/97 permiten desarrollar un modelo con dos niveles de análisis: **el alumno** (nivel 1) y **la escuela** (nivel 2)⁴.

Test estadístico. Para decidir si la asociación es estadísticamente significativa se usa el test de la razón de máxima verosimilitud. Bajo hipótesis de nulidad de diferencia igual a 0 (cero), la diferencia entre valores de máxima verosimilitud de dos modelos sigue la distribución de chi-cuadrado con grados de libertad (gl.) igual al número de nuevos parámetros. En general, la evaluación se realiza por cada factor individual. La probabilidad de ocurrencia es indicada en cada caso.

⁴ Casi en la totalidad de la muestra de la ECE/97 fue sorteado un solo grupo (sección) por escuela. En este caso, y a los fines del análisis, ambos términos son sinónimos y expresan el nivel 2 de agregación. Obsérvese que si se contase con información de varias secciones por escuela, los alumnos serían el nivel 1, los grupos (secciones) el nivel 2 y las escuelas el nivel 3.

Capítulo 2

LA DISTRIBUCIÓN Y LAS DESIGUALDADES DEL RENDIMIENTO ;Error! Marcador no definido. EN LENGUA Y MATEMÁTICA⁵

No todos los alumnos de 7º año obtuvieron el mismo puntaje en las pruebas de Matemática y Lengua. Por el contrario, algunos alumnos lograron puntajes más altos o más bajos que otros. Por otro lado, las escuelas también difieren entre sí respecto al rendimiento promedio que alcanzaron sus alumnos. Es decir, el rendimiento varía y la magnitud total de esa variación es el resultado de dos componentes:

. la variación del rendimiento promedio de cada escuela en torno al rendimiento promedio de todas las escuelas y

. la variación del rendimiento de cada alumno en torno al rendimiento promedio de su escuela (o sección).

De aquí se infiere que el rendimiento de un alumno se diferencia de (se parece a) otro por sus distancias diferentes (similares) respecto al rendimiento promedio de **su** escuela y al rendimiento promedio general de **todas** las escuelas.

En este capítulo nos dedicaremos a determinar la magnitud de ambas variaciones y su grado de significación estadística. Pero previamente, parece necesario explicar por qué se considera pertinente e importante realizar esta tarea. Comenzaremos revisando algunos conceptos básicos.

Es razonable pensar que una proporción de las diferencias de rendimiento entre alumnos y entre escuelas se deba a errores en la medición o a causas fortuitas. Por ejemplo, es posible que un alumno con un bajo nivel de aprendizaje haya conseguido igual puntaje en la prueba que otro con un nivel de aprendizaje más alto, simplemente porque fue hábil para "copiar" las respuestas a su compañero más próximo; o porque, el día de la prueba, el mejor alumno tenía un fuerte dolor de cabeza que lo indujo a errores que no hubiera cometido en otra situación; o porque el alumno eligió las respuestas al azar y acertó en todas ellas, aunque esto sea muy poco probable.

Este tipo de situaciones, sin embargo, no pueden explicar sino una parte pequeña de la variación del rendimiento. Si así no fuese, podríamos dudar de la calidad (confiabilidad) de la evaluación. La proporción más extensa de las desigualdades en el rendimiento seguramente no es aleatoria sino que puede ser explicada por determinadas características. Para que ello sea posible, tales características deben también variar de alumno en alumno o de escuela en escuela. Por eso se las llama "variables".

⁵ En este Capítulo, además de Matemática, se analizan también los datos de Lengua.. Los capítulos subsecuentes se refieren exclusivamente al área de Matemática.

* Así por ejemplo, la edad del alumno varía entre los alumnos, es decir, no todos los alumnos tienen la misma edad; la disponibilidad de recursos didácticos no es la misma en todas las escuelas.

Algunas de las variables capaces de explicar las diferencias en el rendimiento nunca serán conocidas ni mensuradas; a otras en cambio, las podremos identificar y medir. Si analizamos el comportamiento de cualquiera de estas últimas en relación con el rendimiento, podremos verificar si las variaciones de ambas se asocian, es decir, si existe una variación concomitante entre ambas y en qué sentido o dirección.

* Por ejemplo, si a medida que aumenta la edad del alumno, disminuye el rendimiento (sentido negativo); o si a medida que las escuelas tienen más recursos didácticos, el rendimiento de sus alumnos es más alto (sentido positivo).

En este caso, decimos que las variables nos ayudan a "explicar" las variaciones en el rendimiento y son comúnmente denominadas **factores del rendimiento**.

Tradicionalmente, esos factores han sido clasificados en dos categorías - factores escolares y extra-escolares- que de hecho, suponen la distinción de los dos componentes de la variación del rendimiento, tal cual lo expusimos anteriormente, es decir, las desigualdades "entre-escuela" y "entre-alumno".

. **Los factores escolares.** Es obvio que las diferencias entre los rendimientos promedios de las escuelas se deben explicar principalmente por características que son comunes a todos los alumnos de una misma escuela, o sea características grupales.

* Ejemplo: los alumnos pertenecientes a un mismo curso participan homogéneamente de algunas características (por ejemplo, el maestro emplea una metodología activa de enseñanza) y simultáneamente, todos ellos se diferencian por igual de los alumnos de otra sección (con un maestro que aplica un método "frontal" pasivo). La tarea relevante es demostrar, por ejemplo, que la diferencia entre sus rendimientos promedios respectivos se debe al método de enseñanza aplicado por el maestro.

. **Factores extra-escolares.** Por otro lado, las distancias entre los rendimientos obtenidos por los alumnos de una misma escuela (o curso) podrán explicarse por las diferencias en sus características personales.

* Ejemplo: En una misma sección, algunos alumnos serán repitentes y otros no; algunos tendrán más motivación que otros para las actividades escolares; algunos provendrán de un medio familiar socialmente más aventajado que otros, etc. En este caso, sería de interés saber si las diferencias en los rendimientos al interior de esa sección se deben, en parte, a algunas de esas características.

Estamos en condiciones ahora de argumentar sobre la importancia de especificar la distribución de la variación del rendimiento. Sin dudas, podemos asumir que toda política educativa y la correspondiente implementación programática, pretende afectar la distribución de los **factores escolares** con la finalidad de mejorar el aprendizaje de los alumnos. Dimensionar la importancia

relativa de los factores escolares y extra-escolares en la determinación del aprendizaje es de gran interés y utilidad porque permite saber cuánto se puede llegar a mejorar la calidad de la educación a través de esas políticas. De hecho, estaremos dando respuesta a una pregunta clave en el área educativa, a saber: **¿La escuela afecta el aprendizaje del alumno?, ¿O éste sólo refleja los condicionantes externos a la escuela? Si ese efecto existe, ¿es estadísticamente significativo?** Si relacionamos estos interrogantes con la distinción de los componentes de la variación del rendimiento, según fue expuesto anteriormente, percibiremos que la forma operacional más adecuada para resumir aquellas preguntas es la siguiente: **¿la variación del rendimiento promedio de las escuelas es estadísticamente significativo?**

Objetivos específicos. La metodología utilizada en este trabajo para el análisis de los datos (ver Capítulo 1) permite determinar justamente qué proporción de las desigualdades en el rendimiento se explica por factores que operan a través del agrupamiento escolar (diferencias entre escuelas) y cuál por factores que operan a nivel individual (diferencias entre alumnos). Con ella, podremos responder a las siguientes preguntas:

1. ¿Cuál es el peso relativo que tienen los factores escolares y extra-escolares en la determinación del rendimiento de Matemática y Lengua? O sea, ¿cómo se distribuye la variación de los rendimientos escolares (en Lengua y Matemática) entre **escuelas** y **alumnos**?

* En términos más técnicos: ¿cómo se descompone la varianza total del rendimiento (en Lengua y Matemática) entre sus dos componentes: la varianza "**entre-escuela**" y la varianza "**intra-escuela**" (o "entre-alumno")?

2. ¿La magnitud de la variación del rendimiento promedio de las escuelas es importante? Si aceptamos al rendimiento como un "proxy" adecuado de la calidad educativa, la pregunta es: ¿Las escuelas, difieren entre sí respecto a la **calidad educativa** que alcanzan?

* En otras palabras, ¿la estimación de la varianza "entre-escuela" del rendimiento promedio es significativamente distante de su error standard?

Metodología de análisis. Tal cual fuera expuesto en el Capítulo 1, para el análisis de los datos se utiliza la técnica de niveles múltiples. Allí se explicitaron algunos de sus conceptos básicos. De aquí en adelante, cada vez que realicemos un nuevo tipo de operación, expondremos brevemente su significado.

Para lograr el objetivo planteado en este capítulo, realizaremos la operación inicial de la técnica de análisis, que incluye dos momentos, a saber: (a) la determinación de la variación total del rendimiento alrededor de su media global; y (b) la partición de esa variación total en sus dos componentes: la variación "entre-alumno" y la variación "entre-escuela". La particularidad de esta técnica es que estima la variación alrededor de la media global del rendimiento involucrando **simultáneamente** a los dos niveles (alumnos y escuelas). Nos permite saber entonces, cómo varían los promedios de las escuelas y los puntajes de los alumnos en torno a esa media.

* Para calcular ese parámetro y las variaciones "entre-escuela" y "entre-alumno", se regresan los rendimientos escolares de cada alumno ($REND_{ij}$) sobre una constante (cons) que asume un valor = 1 para todos los alumnos. Formalmente la operación se expresa así:

$$\begin{aligned} \text{rend}_{ij} &\sim N(XB, \Omega) \\ \text{rend}_{ij} &= \beta_{0ij} \text{cons} \\ \beta_{0ij} &= \beta_0 + u_{0j} + e_{0ij} \\ [u_{0j}] &\sim N(0, \Omega_u) : \Omega_u = \begin{bmatrix} \sigma_u^2 & 0 \\ 0 & 0 \end{bmatrix} \\ [e_{0ij}] &\sim N(0, \Omega_e) : \Omega_e = \begin{bmatrix} \sigma_e^2 & 0 \\ 0 & 0 \end{bmatrix} \end{aligned}$$

, donde $(\text{rend})_{ij}$ es el rendimiento del alumno i en la escuela j ; (cons) es una constante con valor 1; β_0 es el rendimiento promedio estimado; e_{ij} es el "residuo" a nivel alumno, para el i -ésimo alumno en la j -ésima escuela y u_j es el residuo de nivel 2 (escuela). Las últimas dos (u_j y e_{ij}) son cantidades **aleatorias**, con una media=0, no correlacionadas y con distribución normal. Por lo tanto, podemos estimar sus varianzas, indicadas como σ_u^2 y σ_e^2 , respectivamente, quedando especificado así el porcentaje de la variación total del rendimiento que se debe a las diferencias entre las escuelas (correlación "intra-escuela"⁶). Ella representa principalmente el peso que tienen las características grupales en la explicación de las variaciones totales de los rendimientos escolares.

Al conjunto de estas estimaciones se las denomina modelo "vacío" porque no contiene ningún predictor del rendimiento, sino solamente los niveles de agregación intervinientes (alumno y escuela). La media global se encuentra en la parte **fija** del modelo y las estimaciones de las variaciones "entre-escuela" y "entre-alumno" forman la parte **aleatoria** del modelo.

El modelo "vacío" es el punto de partida del análisis porque define una "línea de base" que sirve como patrón de referencia comparativa. El análisis que se realiza en el próximo capítulo permitirá entender completamente esta idea.

En base a los datos disponibles, existen dos formas de especificar el segundo nivel de agregación: la escuela o la sección. Nos parece pertinente evaluar las posibles diferencias que puedan surgir en los resultados como consecuencia de elegir uno cualquiera de ellos.

Resultados. Comentaremos ahora los resultados obtenidos en Matemática y Lengua, usando esos dos criterios de agregación. Además, recordemos que las secciones o escuelas deben tener al menos 20 alumnos con información válida para ser incluidas en todos los análisis presentados en este trabajo (ver Capítulo 1). Con esta restricción, los tamaños de las muestras quedaron definidos así:

	Matemática	Lengua
Escuelas	890	820
Alumnos	23.895	21.864
Secciones	911	832

⁶ Para el caso de un modelo de 2 niveles, donde sección y escuela coinciden, la correlación "intra-escuela" mide la proporción de la varianza total debida a la variación entre escuelas .

Alumnos	23.013	20.769
---------	--------	--------

Los principales resultados obtenidos con el archivo de datos donde el nivel 2 de agregación es la escuela, son los siguientes:

2.1 Modelo vacío - Lengua		2.2 Modelo vacío Matemática	
58,041 (0,416)		49,848 (0,457)	
CONS	131,825 (6,992)	CONS	177,336 (8,812)
CONS	247,344 (2,411)	CONS	218,657 (2,039)
Test de máxima verosimilitud: 184744,9		Test de máxima verosimilitud: 199302,200	

El rendimiento promedio global es estimado en 58,041% para Lengua y 49,848% para Matemática⁷ (Cuadro 2.1 y 2.2). Es decir, el rendimiento en Matemática es inferior al de Lengua en 10 puntos porcentuales.

. La varianza "entre-escuela" de los puntajes promedios de las escuelas es estimada en 131.825 en Lengua y 177.336 en Matemática. Dado que la magnitud de los errores standard (6,992 y 8,812, respectivamente) es extremadamente menor que la estimación correspondiente, los datos indican que las diferencias de rendimiento promedio entre las escuelas es altamente significativa. Si aceptamos que el rendimiento promedio de las escuelas indica calidad educativa (como producto), entonces, **las escuelas se diferencian significativamente entre sí respecto a la calidad educativa que alcanzan.**

. Por otro lado, las estimaciones mencionadas significan que 34,8% de la variación total de los rendimientos en Lengua y 44,8% en Matemática se debe a diferencias entre las escuelas;

* Estos porcentajes surgen de relacionar la variación estimada para el nivel 2 con la variación total (Lengua=131,825+247,344; Matemática=177,336+218,65). A esta proporción de la varianza total se la denomina correlación "intra-escuela".

Dado que se trata de la variación de los promedios de las escuelas en torno del rendimiento promedio general, esos porcentajes representan el peso relativo que tienen los factores escolares (características de la escuela, en sentido amplio) en la explicación de la variación del rendimiento. Es obvio que quienes pueden explicar las diferencias en el rendimiento promedio de las escuelas son principalmente, aunque no totalmente, diferencias entre las propias escuelas respecto a

⁷ Estas estimaciones no deben coincidir necesariamente con las divulgadas por el MCyE, dada las diferencias en las bases de datos utilizadas. Las estimaciones del MCyE se basan en la totalidad de la muestra, mientras que las que estamos presentando se extraen de una base de datos que cumple con las restricciones anteriormente expuestas .

determinadas características institucionales. Por eso, la variación detectada puede ser entendida como una consecuencia de la selectividad del sistema educativo a través del agrupamiento por escuelas.

. Respecto a la comparación entre ambas áreas de conocimiento, el mayor porcentaje en Matemática indica que el aprendizaje de esta disciplina es más heterogéneo en el agrupamiento escolar, o sea, el efecto de la selección o entrada en una determinada escuela es menos pronunciado en Lengua que en Matemática. En consecuencia, **las desigualdades del rendimiento en Matemática están más vinculadas a factores estrictamente escolares**, ofreciendo por lo tanto, un espacio más promisorio para la búsqueda de factores propios del sistema escolar que las expliquen.

. Finalmente, la variación "entre-alumno" constituye el 59.2% de la variación total en Lengua y 47,2% en Matemática. En este caso, la explicación debe buscarse entre los factores vinculados al propio alumno y no al agrupamiento en el sistema educativo.

En resumen, los resultados permiten concluir, aunque provisoriamente, que pertenecer a una escuela determinada implica una mayor (menor) probabilidad de acceder a un nivel de calidad educativa (como producto) significativamente más alta (baja) que la esperada en otras escuelas, distancia explicada en gran medida, por las diferencias institucionales entre ellas (factores escolares).

Escuelas y secciones. Los mismos procesamientos cuyos resultados hemos analizados hasta aquí fueron repetidos, pero ahora usando como unidad de análisis a las secciones y no a las escuelas. Los resultados se observan en los Cuadros 2.3 y 2.4. Si los comparamos con los anteriores, observamos que no existen diferencias significativas. Es decir, las conclusiones a las que hemos arribado anteriormente son también válidas cuando usamos a la sección como nivel 2 de agregación.

* Esta conclusión nos ayuda también a tomar la decisión acerca de cuál agregación (escuela o sección) usar en el resto del análisis. Se torna razonable que cuando cualquiera de los dos es elegible, prefiramos trabajar con el dato agregado a nivel de escuela; en algunos casos, sin embargo, deberemos desarrollar el análisis tomando la sección como criterio de agregación.

2.3 Modelo vacío Lengua (sección)	
CONS	
58,132	
(0,419)	
☺	CONS
CONS	136,594
	(7,139)
☺	CONS
CONS	239,431
	(2,402)
Test de máxima verosimilitud: 174975,2	

2.4 Modelo vacío Matemática (sección)	
CONS	
49,879	
(0,454)	
☺	CONS
CONS	179,333
	(8,811)
☺	CONS
CONS	214,999
	(2,045)
Test de máxima verosimilitud: 191710,3	

Las estimaciones anuales.

¿Cómo se sitúan los resultados presentados hasta aquí, correspondientes al 7º año de la EGB de 1997, respecto a otros, relativos a otros años y niveles? En el Cuadro 2.5 se presentan las estimaciones anuales para 7º y 3º año de la EGB, desde 1994.

Debemos ser muy precavidos a la hora de extraer conclusiones de estos datos. En primer lugar, observamos que las estimaciones oscilan de año en año. Una parte de estas oscilaciones seguramente deben ser imputadas a cambios en las pruebas aplicadas, en las condiciones concretas de su aplicación y en las características de la muestra "real" obtenida. Por lo tanto, no es recomendable inferir cualquier conclusión acerca de "la tendencia temporal". Afirmaciones tales como "la importancia de los factores escolares ha disminuido (o aumentado) durante este período" deben ser sencillamente dejadas de lado.

En segundo lugar, los resultados de 7º y 3º se refieren a pruebas que difieren en cuanto a su extensión y a las "competencias" medidas, y estas diferencias pueden también afectar las estimaciones de correlación "intraclase".

Finalmente, recordemos que cuando comparamos estimaciones extraídas de muestras, debemos considerar el "intervalo de confianza" en torno a la estimación puntual, determinado por el error standard, lo cual implica que dos estimaciones no son diferentes si ambas tienen al menos un punto de sus intervalos en común.

Conscientes de estas restricciones, pasemos a leer las informaciones presentadas:

1. En primer lugar, todas las estimaciones son convergentes con la conclusión más relevante a la que hemos arribado anteriormente: los factores escolares, es decir, las diferencias entre las escuelas, tienen un gran peso en la explicación de la variación total en el aprendizaje de los alumnos;
2. También se confirma que el aprendizaje escolar de la matemática de 7º grado está más vinculado a los factores del agrupamiento escolar que el de la Lengua. La escuela produce fuertes diferencias en el acceso al conocimiento de la Matemática;
3. Esta última conclusión no es válida cuando se trata de los alumnos de tercer año. En ambas disciplinas, el peso de los factores escolares es prácticamente el mismo. Los datos parecen reflejar el efecto acumulativo del sistema educativo. Próximo al ingreso en el sistema, los niveles de conocimiento exigidos en ambas disciplinas son "básicos" y por tanto, ambos se encuentran a la misma distancia respecto al que existe en el entorno familiar de origen. Esto conduce a que el efecto de las variaciones socioculturales de las familias u otras características personales del alumno sean más similares en ambas áreas del conocimiento. Al pasar el tiempo, a medida que se avanza en el sistema, la escuela produce un efecto mayor sobre el aprendizaje de aquella área un poco más alejada del saber familiar y más próxima a un saber producido por la propia escuela.

2.5 Porcentaje de la varianza total debida a la variación "entre-escuela" 3º y 7º año EGB - Matemática y Lengua - 1994 - 1997.

	7º año		3º año	
	Matem.	Lengua	Matema.	Lengua
1994	44,4	35,4	-	-
1995	52,8	40,8	42,6	44,1
1996	51,9	41,8	36,7	40,0
1997	44,8	34,8	46,6	47,0

Capítulo 3

LA EQUIDAD EDUCATIVA

En el capítulo anterior hemos especificado los dos componentes de la variación del rendimiento ("entre-alumno" y "entre-escuela"), sugiriendo sus vínculos con el concepto de calidad educativa. Si aceptamos al rendimiento, medido con una prueba de opciones múltiples, como un indicador adecuado, aunque no libre de críticas, de la "calidad educativa" (como producto), obtenemos dos definiciones **provisorias**: una se refiere al rendimiento del alumno individual en relación a sus colegas de escuela o sección; la otra nos remite al **rendimiento promedio obtenido por los alumnos de una escuela (o sección), en relación al de otras escuela**. A esta última se la puede denominar **calidad educativa institucional**.

También introdujimos la idea de que al menos una parte de las desigualdades de rendimiento entre alumnos y entre escuelas, puedan ser explicadas por determinadas situaciones o variables. Para comprobar esta hipótesis, debemos analizar el comportamiento de cualquiera de esas variables en relación con el rendimiento, verificando si las variaciones de ambas se asocian, es decir, si existe una variación concomitante entre ambas y en qué sentido o dirección. Finalmente, hemos afirmado que aquellas variables que muestran una estrecha asociación con el rendimiento se denominan "factores del rendimiento", siendo posible distinguir los escolares y los extra-escolares.

Nos interesa ahora la relación entre el rendimiento del alumno y su origen social. El sentido común y numerosos antecedentes de investigación, inducen a hipotetizar como comportamiento más probable, que el rendimiento mejorará a medida que ascendemos en el nivel socioeconómico del alumno (variable individual). También es razonable esperar que el rendimiento promedio (de la escuela) varíe en función de la composición social de la escuela (variable grupal) y en aquel mismo sentido. Diremos, entonces, que esa variable (sea individual o contextual) se asocia positivamente (+) con el rendimiento.

Con base en este conjunto de ideas, podemos definir operacionalmente a la **equidad educativa como la magnitud o intensidad de la asociación entre rendimiento y origen socioeconómico del alumno**⁸. Si nos situamos a nivel institucional, es decir, a nivel de escuela, este concepto invoca la capacidad que tiene una escuela (o sección) para compensar el efecto que tiene el origen socioeconómico del alumno sobre su aprendizaje, es decir, los rendimientos obtenidos por los alumnos en las escuelas más equitativas dependen menos del origen socioeconómico del alumno que en las escuelas menos equitativas.

⁸ Adviértase que la equidad puede predicarse respecto a diferentes variables. El nivel socioeconómico podría reemplazarse por "sexo", "disponibilidad de recursos didácticos", "repitencia", etc. Existe obviamente, una excepción. Más allá de la polémica conceptual y metodológica que la rodea, no sería razonable incluir a la "inteligencia" o a la "aptitud" entre esas variables.

Antes de continuar, debemos volver brevemente al concepto provisorio de calidad propuesto anteriormente. Ahora sabemos que esa definición debe incorporar explícitamente el hecho de que el rendimiento varíe en función del nivel socioeconómico. Entonces, debemos entender la **calidad educativa institucional como el rendimiento promedio obtenido en la prueba por los alumnos de una escuela, ajustado por el origen socioeconómico de esos alumnos**. Observemos que esta definición limita o condiciona el análisis de los datos. Así por ejemplo, se pueden comparar los rendimientos promedios de dos escuelas, sólo cuando la composición social de ambas sea similar.

Es fácil percibir, siempre desde el punto de vista de la escuela (nivel 2 de agregación), que ambos conceptos, calidad y equidad, están indisolublemente ligados. En realidad, se los puede considerar como dimensiones de un único concepto, el de **eficacia institucional**: escuelas que **simultáneamente consiguen**,

- (i) un rendimiento promedio más alto que el esperado en base a la composición social de su alumnado (**calidad**), y
- (ii) una mayor compensación de los efectos del origen socioeconómico del alumno sobre su rendimiento (**equidad**).

Es decir, **eficaces** son aquellas escuelas capaces de obtener **simultáneamente mayor calidad** (la excelencia en los aprendizajes) **y equidad** (distribución social del conocimiento).

* Ejemplo 1: De dos escuelas cuyos alumnos tienen un origen socioeconómico similar, es más eficaz aquella cuyos alumnos consiguen un rendimiento promedio más alto;

* Ejemplo 2: Si la escuela A con alumnos de origen socioeconómico más desfavorable que la escuela B, consigue un rendimiento promedio igual o mayor que el de B, entonces, la escuela A es más eficaz que la B.

El aspecto más importante de este razonamiento es que considera a la calidad y a la equidad institucional como variables a ser estudiadas. El foco de atención se centra en la desigualdad (variación) institucional ("entre-escuelas") de la calidad y equidad educativa. En el capítulo anterior hemos iniciado el análisis de la calidad. En éste expondremos los resultados obtenidos con una estrategia de análisis que permite focalizar la calidad y la equidad tal cual han sido definidas anteriormente.

Objetivos. En términos generales, los interrogantes a responder son:

1. ¿El sistema educativo es **equitativo**?
2. ¿El grado de **equidad educativa** varía entre las escuelas del sistema?
3. ¿Las escuelas difieren respecto a la **calidad educativa** que alcanzan?

Un objetivo instrumental. Además de estos objetivos conceptuales, nos proponemos otro de carácter más instrumental, aunque imprescindible para poder desarrollar el análisis presentado en los capítulos subsiguientes. En base a diferentes indicadores del origen social del alumno, trataremos de **identificar el modelo más eficiente para representar las relaciones entre ese factor y el rendimiento del alumno**.

* En este contexto particular, por **modelo** entendemos la representación (sistema de ecuaciones) del efecto de una o varias variables (independientes) sobre el rendimiento (variable dependiente).

* La **eficiencia** del modelo depende de la relación entre su capacidad predictiva y el número de variables que lo componen. De dos modelos que expliquen una misma proporción de la variación del rendimiento, será mejor valorado aquel que contenga el menor número de variables (parsimonia).

Desde 1993, los cuestionarios aplicados a los alumnos de 7º año de la E.G.B. han contemplado la medición de su origen social. Dado que los cuestionarios han sido del tipo "auto-aplicado", siempre se optó por medir esa característica a través de la educación del padre y de la madre, y de la posesión familiar de determinados bienes de uso durable. Con base en esas mediciones, también ha sido posible construir la variable "contextual" correspondiente, es decir, el nivel socioeconómico promedio de la escuela o del aula ("composición social").

En el cuestionario de 1997⁹ se incluyeron algunas preguntas adicionales, destinadas a fortalecer aquella medición. Con base en ellas, es posible construir una nueva medición individual, "el grado de acceso a bienes culturales", y la contextual correspondiente, que denominaremos "**contexto cultural**" de la escuela.

Estrategia metodológica. Continuaremos usando la técnica de análisis estadístico por niveles múltiples, introducida en el capítulo anterior. El procedimiento para responder a las preguntas planteadas inicialmente como objetivos de este capítulo, está organizado en 6 etapas:

1. Estimación de los parámetros del modelo "vacío";
2. Análisis de la asociación entre rendimiento y cada indicador individual (a nivel de alumno);
3. Análisis de modelos que combinan los indicadores individuales;
4. Análisis de las relaciones entre rendimiento y cada indicador grupal (a nivel de escuela);
5. Análisis de los modelos que combinan indicadores individuales y grupales.
6. Evaluación de la variación de la calidad y equidad institucional.

* En los cinco primeros pasos nos limitaremos a analizar la "parte fija" de los modelos (ver explicación más adelante), mientras que en el último, el centro de atención es su "parte aleatoria".

Los indicadores. Las mediciones que serán consideradas son las siguientes:

- ~~??~~ (matema): % de respuestas correctas en la prueba de Matemática;
~~??~~(Ibie4): Índice sumativo de tenencia de 17 bienes de uso durables en el hogar ;
?? (Iedu4): Índice sumativo de la educación del padre y de la madre;
?? (NSE): Nivel socioeconómico familiar: índice sumativo de (Ibie4) y (Iedu4)¹⁰;
?? (NSE_MED) (NSE) promedio en la escuela;
?? (Icul2r2): Índice sumativo de la posesión de bienes culturales;
?? (CULT_MED): (Icul2r2) promedio en la escuela.

⁹ Ver en Anexo un ejemplar del Cuestionario del alumno de 1997.

¹⁰ La metodología empleada para la creación de este índice se encuentra en "La medición del nivel socioeconómico familiar en el nivel primario –Una opción metodológica y sus resultados- 1994/1995/1996/1997 ", Rubén Cervini y otros, en **Informe de Investigación** N°48. Ministerio de Cultura y Educación. Buenos Aires. 1998.

La muestra. El archivo de datos analizados es de 25.271 alumnos y 995 secciones.

Test de significación. Para evaluar el grado de asociación con el rendimiento se utiliza el test de máxima verosimilitud, tal cual fuera explicado en el Capítulo 1.

* Una forma visualmente directa de percibir la asociación es observando la disminución relativa que cada indicador o conjunto de ellos produce sobre la variación no-explicada del rendimiento. En otras palabras, el criterio de eficacia nos induce a observar la proporción de varianza del rendimiento que es explicada por el indicador bajo análisis.

Resultados. En el Cuadro 3.1 se presenta una síntesis de los resultados obtenidos en cada uno de los cinco primeros pasos.

1. El modelo "vacío". La primera operación consiste en estimar los parámetros del modelo "vacío", es decir, la partición de la varianza "entre-escuela" y "entre-alumno". Los valores relativos de las estimaciones obtenidas no difieren de las registradas en el capítulo anterior (ver Cuadro 2.2 del Cap. 2), lo cual nos exime de su análisis. Debemos retener, sin embargo, el valor arrojado por el test de máxima verosimilitud (=210501,100), ya que constituye el punto de partida y el valor de referencia para evaluar los resultados que se obtendrán con los próximos modelos en este capítulo.

2. Mediciones individuales. El objetivo es evaluar el grado de asociación inicial de cada medición con el rendimiento. Haremos una breve exposición de los conceptos implicados en esta operación, ejemplificando con la variable "nivel socioeconómico del alumno" (NSE). La misma lógica será aplicada a las variables restantes.

Cuadro 3.1 - Parámetros estimados para los diferentes indicadores del origen social del alumno y de la composición sociocultural de la escuela.

Indicadores	Test de verosimilitud	Razón de verosimilitud	Variación nivel 2 (escuela)	Disminución variación nivel 2
1. Modelo "vacío"	210501,100	-	180,109	-
2. Mediciones individuales				
(1) NSE	10020,300	480,8	157,671	12,46
(2) Ibie4	210252,800	248,3	165,832	7,92
(3) Iedu4	210110,700	390,4	163,936	8,97
(4) Icul2r2	210044,000	457,1	164,205	8,83
3. Múltiples mediciones individuales				
(2) + (3)	210009,000	492,1	157,879	12,34
(3) + (4)	209835,200	665,9	156,093	13,33
(1) + (4)	209818,200	682,9	153,297	14,89
4. Mediciones grupales				
(5) NSE_MED	210160,100	341,0	125,298	30,43
(6) CULT_MED	210196,400	304,7	130,310	27,60
5. Mediciones individuales y grupales				
(5) + (1)	209798,600	702,5	125,424	30,36
(6) + (4)	209823,900	677,2	130,440	27,50
(1)+(4)+(5)	209615,800	885,3	124,518	30,86
(1)+(4)+(5)+(6)	209609,600	891,5	123,688	31,33

Paréntesis metodológico 1

Inicialmente, podríamos afirmar que el objetivo propuesto consiste en determinar la proporción de la variación (varianza) total del rendimiento en Matemática (REND) que es "explicada" por (NSE).

* Tradicionalmente, cuando se quiere conocer la relación entre el puntaje obtenido en una prueba (y_i :REND) y el nivel socioeconómico (x_i :NSE) se aplica la técnica bien conocida de regresión múltiple ordinaria, a través de la cual se estima una ecuación que expresa el rendimiento como una función lineal de (NSE):

$$(REND)_i = a + b(NSE)_i + e_i \quad (1)$$

donde i toma valores diferentes para cada alumno; a es el "intercepto" y b es la pendiente, ambos parámetros a ser estimados; e_i es el residuo, al que se lo supone con media=0 y varianza constante; $(REND)_i$ y $(NSE)_i$ son los valores del rendimiento y del nivel socio económico para el alumno i .

No obstante, con este modelo no es posible distinguir las variaciones "entre-alumno" y "entre-escuela" de (REND). Se asume que dos alumnos situados en el mismo (NSE) deberán tener los mismos puntajes, sin importar la(s) escuela(s) a la(s) que asistan. Pero, en realidad, es probable que la pendiente b varíe de escuela en escuela y parece importante determinar si es así. Aquí radica la principal diferencia entre el análisis de niveles múltiples con el análisis de covarianza. Esta última no permite generalizar la aleatoriedad a diversos coeficientes en varios niveles, y de esta forma, utilizar tests de significación para identificar los factores que, incorporados al modelo, reduzcan la variabilidad entre alumnos, grupos o escuelas.

Sin embargo, nuestro interés va más allá. Por las razones conceptuales expuestas hasta aquí, necesitamos saber también qué sucede con la variación del rendimiento "entre-alumno" y "entre-escuela". La operación para responder a esta curiosidad consiste en controlar (NSE) y analizar los efectos sobre (REND) en los dos niveles (alumno y escuela). Para ello, se introduce (NSE) en la parte fija del modelo "vacío", obtenido en el paso anterior, y se estiman nuevamente los parámetros.

?? Si se pretendiese estudiar una sola escuela, el modelo en la ecuación (1) sería suficiente, ya que la variación de e_i (residuo del nivel 1) sería la variación total y significaría la parte de los rendimientos observados no "explicados" por el modelo. Pero, en nuestro caso, debemos extraer conclusiones de numerosas escuelas que son una muestra del universo de escuelas en el país. En este caso, el modelo completo se expresa como:

$$(REND)_{ij} = a_j + b(NSE)_{ij} + e_{ij} \quad (2)$$

donde $(\text{REND})_{ij}$ y $(\text{NSE})_{ij}$ son el rendimiento y el nivel socioeconómico del alumno i en la escuela j , a_j es el intercepto de la regresión para la j -ésima escuela, b es la pendiente de la regresión que, por el momento, se supone constante en todas las escuelas y e_{ij} es el residuo a nivel alumno, para el i -ésimo alumno en la j -ésima escuela.

Por otro lado, a_j tiene dos componentes: el intercepto de la predicción para todas las escuelas, a , y la diferencia entre el intercepto observado en la j -ésima escuela y a ($?_j = a_j - a$), o sea, el residuo de nivel 2 (escuela). Podemos re-escribir (2) como:

$$(\text{REND})_{ij} = a + b(\text{NSE})_{ij} + ?_j + e_{ij} \quad (3)$$

Aquí, $?_j$ y e_{ij} son cantidades **aleatorias**, con una media=0, no correlacionadas y con distribución normal. Por lo tanto, sólo debemos estimar sus varianzas, $?_?$ y $?_e$. Por otro lado, a y b son **fijas**, comunes a toda la población, y también serán estimadas. La existencia de dos parámetros aleatorios en la ecuación (3) es precisamente lo que caracteriza un modelo de niveles múltiples. Por lo tanto, deberán estimarse 4 parámetros. Para ello se procede de la siguiente manera. El parámetro **fijo** b es asociado con (NSE) mientras que los otros tres son asociados con la constante (CONS) anteriormente utilizada (ver paso anterior y Capítulo 1). De manera que la ecuación en (3) se especifica como

$$(\text{REND})_{ij} = a(\text{CONS}) + b(\text{NSE})_{ij} + ?_j(\text{CONS}) + e_{ij}(\text{CONS}) \quad (4)$$

Con el objetivo de diferenciar visualmente las estimaciones de la parte fija y de la parte aleatoria, la ecuación en (4) se puede re-escribir:

$$\begin{aligned} \text{rend}_{ij} &\sim N(XB, \Omega) \\ \text{rend}_{ij} &= \beta_{0ij} \text{cons} + \beta_1 \text{nse}_{ij} \\ \beta_{0ij} &= \beta_0 + \mu_{0ij} + e_{0ij} \end{aligned} \quad (5)$$

$$\begin{aligned} \begin{bmatrix} \mu_{0ij} \end{bmatrix} &\sim N(0, \Omega_\mu) : \Omega_\mu = \begin{bmatrix} \sigma_\mu^2 & 0 \\ 0 & 0 \end{bmatrix} \\ \begin{bmatrix} e_{0ij} \end{bmatrix} &\sim N(0, \Omega_e) : \Omega_e = \begin{bmatrix} \sigma_e^2 & 0 \\ 0 & 0 \end{bmatrix} \end{aligned}$$

Obsérvese que β_{0ij} incluye el parámetro fijo estimado (β_0) rendimiento promedio) y la estimación de su variación aleatoria a nivel de escuela ($?_j$) y de alumno (e_{ij}).

Recordemos que la **parte fija** del modelo se refiere a los parámetros que permiten definir una línea promedio para **todos** lo alumnos de **todas** las escuelas. Se supone que la intensidad de la asociación entre (NSE) y (REND) es constante para todas las escuelas del sistema. Ello es así porque, en nuestro ejemplo, en la **parte fija** del modelo se estiman:

. el valor promedio de (REND), ajustado por (NSE); ese valor está indicado por el punto donde la línea que representa la relación entre ambas variables corta a la ordenada (rendimiento), y

. el valor promedio de la fuerza de la relación entre (REND) Y (NSE), es decir, la pendiente de la línea que representa la relación entre ambas variables.

De acuerdo a los conceptos expuestos anteriormente, podemos afirmar que en **esta parte del modelo se estiman tanto el nivel de calidad (β_0) y de equidad (β_1) del sistema educativo como un todo, cuanto la variación de la calidad educativa institucional (σ^2).**

El resultado obtenido al introducir (NSE) en el modelo "vacío" de matemática (matema), es el siguiente:

$$\text{matema}_{ij} \sim N(XB, \Omega)$$

$$\text{matema}_{ij} = \beta_{0ij} \text{cons} + 0,096(0,004) \text{nse}_{ij}$$

$$\beta_{0ij} = 49,829(0,409) + u_{0ij} + e_{0ij}$$

$$\begin{bmatrix} u_{0ij} \end{bmatrix} \sim N(0, \Omega_u) : \Omega_u = \begin{bmatrix} 157,669(7,454) \end{bmatrix}$$

$$\begin{bmatrix} e_{0ij} \end{bmatrix} \sim N(0, \Omega_e) : \Omega_e = \begin{bmatrix} 211,781(1,922) \end{bmatrix}$$

$$-2 * \log(\text{like}) = 210020,200$$

Analizando estos resultados, en la parte fija del modelo podemos observar que la estimación del intercepto (β_0) es 49,829; la pendiente (β_1) es estimada en 0,096. Los errores standard correspondientes (0,409 y 0,004, respectivamente) son muy pequeños comparados con la estimación, indicando la alta significación estadística de tales estimaciones.

* Es decir, la probabilidad de obtener esas estimaciones es muy baja (menor al 1 por mil), si la verdadera correlación fuese igual a cero.

Podemos comprobar esa hipótesis comparando el valor del test de máxima verosimilitud de este modelo (=210020,200) con el del modelo "vacío" (=210501,100). La diferencia es de 480.9 que, con 1 gl. (grado de libertad) tiene una probabilidad de ocurrencia muy inferior al 1 por mil.

Queda definida así la ecuación de la línea general que predice (REND) en base a (NSE):

$$(\text{REND}) = 49,829 + 0,096(\text{NSE})$$

Según propusimos anteriormente, los parámetros de esta ecuación representan el nivel general de **calidad** (rendimiento ajustado por NSE) y **equidad** (la pendiente de la recta) del sistema educativo.

Pero este modelo, a diferencia de la regresión tradicional, ha estimado simultáneamente los dos componentes de la variación del rendimiento: la varianza "entre-escuela" y la varianza "entre-alumno". Esto constituye **la parte aleatoria** del modelo.

La distancia de la línea de cada escuela respecto a la línea general estimada es representada por σ_{0j} . La varianza de esta variable es 157,671. Esta estimación se refiere, entonces, a la variación de las líneas individuales de las escuelas en torno a la línea general estimada (variación en el nivel 2).

* Si se acepta que la distribución de los interceptos de las escuelas es normal, entonces, el 97% de las líneas de las escuelas caerán dentro de la franja definida por $\pm 2 \cdot 157,671$.

La variable e_{0ij} representa la distancia entre el puntaje obtenido por el alumno y el puntaje predicho por la línea de su propia escuela. La variación de esta variable (nivel 1), se estima en 211,780.

En resumen, la variación total del rendimiento ($157,671+211,780=369,451$) ha sido desagregada en sus dos componentes. Ahora bien, si comparamos el valor de la estimación de la variación "entre-escuela" de este modelo (157,671) con el del modelo "vacío" (=180,109), obtenemos una diferencia de 22.438; o sea, la variación "entre-escuela" en el modelo "vacío" disminuyó 12,46%. Esta diferencia expresa la proporción de varianza "entre-escuela" explicada por (NSE). Al mismo tiempo, estamos estimando la magnitud de las diferencias de rendimiento promedio de las escuelas todavía "no explicada", es decir, estamos determinando **el máximo de varianza "entre-escuela" que podría ser explicada por otros factores, después de haber controlado por (NSE).**

Dado que el error standard de esa estimación (7,454) es muy pequeño en relación al valor de la estimación (157,669), se deduce que la variación de los rendimientos promedios de las escuelas **continúa siendo significativa**, aún después de haber controlado (NSE); es decir, a igual nivel socioeconómico, se pueden esperar diferencias pronunciadas entre los rendimientos promedios de las escuelas.

(Fin paréntesis metodológico 1)

Los resultados presentados en la **parte 2 del cuadro 3.1**, permiten extraer las siguientes conclusiones:

a) todas las mediciones mantienen una estrecha relación con el rendimiento;

* La distancia entre el valor de verosimilitud del modelo vacío y el correspondiente a cualquiera de los cuatro términos evaluados tiene una probabilidad de ocurrencia muy por debajo del 1 por mil. Además todas las estimaciones obtenidas fueron varias veces mayores que los errores standard correspondientes.

b) De todas las mediciones experimentadas, (NSE) es la más potente para predecir el rendimiento. El 12,46% de la variación del rendimiento "entre-escuela" se explica por diferencias en el (NSE) de los alumnos. Por lo tanto, esta medición es más eficaz que cualquiera de los dos indicadores que la componen, o sea (Iedu4) y (Icul2r2)

3. Múltiples mediciones individuales. El objetivo de este paso es evaluar el grado de asociación de cada medición con el rendimiento, pero respecto a la variación no explicada por el/los otro/s indicadores. Es decir, tratamos de determinar si cada uno de ellos "explica" una proporción estadísticamente significativa de la variación del rendimiento dejada sin explicar por la/las otra/s variables. Para tal finalidad, debemos incluir simultáneamente más de un indicador en el modelo vacío.

a) Iniciamos el procedimiento considerando la acción conjunta de la disponibilidad de bienes de uso durable en el hogar (Ibie4) y el nivel de instrucción de los padres (Iedu4). Ambos indicadores mantienen su eficacia propia, aún después de haberlos controlado "mutuamente".

* La distancia entre el valor de verosimilitud de este modelo vacío (=210009,0) y el correspondiente al indicador más eficiente de los dos (Iedu4=210110,7) es igual a 101,7 puntos que, con 1g.l., tiene una probabilidad de ocurrencia muy inferior al 1 por mil. Además, las estimaciones obtenidas para ambos términos fueron varias veces mayores que los errores standard correspondientes.

Al comparar estos resultados con los obtenidos individualmente por (NSE) en el paso anterior, vemos que si bien el efecto de este último indicador es menor a la acción conjunta de Ibie4 e Ieduc4, la diferencia no es muy abultada.

* La distancia entre el valor de verosimilitud del modelo con (NSE) (=210020,3) y el que contiene a los 2 términos que estamos analizando (210009,0) es de sólo 11,3 puntos.

Ateniéndonos al criterio de parsimonia, encontramos que es preferible el uso de (NSE) como sustituto de (Ibie4) e (Iedu4); es decir, la decisión de unir los 2 indicadores en uno sólo puede mantenerse.

b) Parece interesante, ahora, evaluar el comportamiento de los dos indicadores más afines: el acceso a bienes culturales (Icul2r2) y el nivel educativo de los padres (Ieduc4). Los resultados

indican que ambos indicadores tienen eficacia predictiva propia sobre el rendimiento, aún cuando se los considere conjuntamente, y por lo tanto, ambos deberían ser tenidos en cuenta si se quiere mejorar la estimación de la relación entre rendimiento y origen social del alumno.

* La distancia entre el valor de verosimilitud del modelo con $Icul2r2$ ($=210044,0$) y el correspondiente al que contiene los dos indicadores ($=209835,2$) es igual a 109 puntos con una probabilidad de ocurrencia muy inferior al 1 por mil. Además, el procesamiento de los datos permitió constatar que ambas estimaciones eran varias veces mayores que los errores standard correspondientes.

c) Finalmente, corresponde evaluar la acción conjunta de (NSE) con el indicador de acceso a bienes culturales. Los resultados son inequívocos: ambas mediciones deben mantenerse. Los resultados indican que casi el 15% de las diferencias entre las escuelas respecto al rendimiento que alcanzan, puede ser explicado por las diferencias en el nivel socioeconómico y en el acceso a bienes culturales de sus alumnos.

* La distancia entre el valor de verosimilitud del modelo individual más eficiente (NSE= $210020,3$) y el correspondiente a la acción conjunta de ambos ($=209818,2$) es igual a 202,1 puntos, con una probabilidad de ocurrencia asociada muy inferior al 1 por mil cuando se tiene 1 g.l. Además se constató que ambas estimaciones son varias veces mayores a los errores standard correspondientes

4. Mediciones grupales. Ahora nuestro objetivo es evaluar el grado de asociación inicial entre cada medición agregada ("contextual") y el rendimiento. Para ello, incluimos cada medición en el modelo vacío una por vez, y observamos los cambios producidos.

a) Ambos indicadores resultan altamente relacionados con el rendimiento. Aún cuando la medición referida a la composición social de la escuela (NSE_MED) es la más significativa, el acceso a bienes culturales (CULT_MED) se sitúa muy próximo a ella: mientras que la primera explica el 30,43% de las diferencias de rendimiento entre las escuelas, CULT_MED lo hace con el 27,60%.

* Los valores de verosimilitud de ambas variables (210160,1 y 210196,4) son significativamente distantes del correspondiente al modelo "vacío" ($=210501,1$).



b) Adviértase también que las dos mediciones consiguen niveles de explicación de la variación "entre-escuela" muy superiores a los obtenidos con los indicadores medidos a nivel del alumno. Estas diferencias se explican justamente por tratarse de mediciones agregadas a nivel de escuela (nivel 2): en general, las diferencias entre escuelas deben explicarse por diferencias en algunas de sus características, por ejemplo, su composición social.

5. Mediciones individuales y grupales. Finalmente, evaluaremos el efecto conjunto de todas las mediciones. Pretendemos conocer el grado de asociación de cada medición con el rendimiento, pero respecto a la variación no explicada por el/los otro/s indicadores. Al igual que en el paso 3, determinaremos si cada indicador "explica" una proporción estadísticamente significativa de la variación del rendimiento dejada sin explicar por la/las otra/s variables. Para tal finalidad, incluiremos simultáneamente más de un indicador en el modelo vacío.

a) Comenzamos observando cómo se comporta la medición individual y la "contextual" correspondiente. Los resultados indican que ambas mediciones deben permanecer, es decir, no sólo el nivel socioeconómico o el acceso a bienes culturales del alumno afectan al rendimiento, sino que la composición social y cultural ayudan también a explicar una proporción significativa de las diferencias entre los rendimientos.



* Las diferencias de los valores de verosimilitud de ambos modelos (209798,6 y 209823,9) respecto a cualquiera de los que se estimaron en los modelos individuales, es superior a 200 puntos, valor con una probabilidad de ocurrencia muy inferior al 1 por mil. Consistentemente, el procesamiento de los datos demostró que, en ambos casos, las estimaciones eran varias veces mayores a los errores standard correspondientes.

b) Habíamos visto que el nivel socioeconómico (NSE) y la disponibilidad de recursos culturales (Icul2r2) mantienen su eficacia predictiva, aún cuando actúen conjuntamente. Esa conclusión continúa siendo cierta, aún cuando se agregue el "control" de la composición social de la escuela (NSE_MED). De hecho, cada una de las tres mediciones explica una proporción significativa de la variación del rendimiento dejada sin explicar por las otras dos. Las estimaciones de este modelo se presentan en el Cuadro 3.2.

3.2 - Modelo con NSE, Icul2r2 y NSE_MED			
CONS	NSE	ICUL2R2	NSE_MED
48,743 (0,372)	0,061 (0,005)	0,065 (0,005)	0,312 (0,021)
	CONS		
CONS	124,517 (5,962)		
	CONS		
CONS	210,193 (1,908)		

El valor de verosimilitud de este modelo (209615,8) es más de 200 puntos superior al del modelo que contiene los dos términos individuales (=209818,2), valor con probabilidad de ocurrencia muy inferior al 1 por mil cuando gl.=1.

c) Finalmente, agregamos al modelo anterior la medición del contexto cultural (Cuadro 3.3).

3.3 - Modelo con NSE, Icul2r2, NSE_MED y CULT_MED				
CONS	NSE	ICUL2R2	NSE_MED	CULT_MED
48,714	0,061	0,064	0,225	0,138
(0,371)	(0,005)	(0,005)	(0,040)	(0,055)
	CONS			
CONS	123,688			
	(5,963)			
	CONS			
CONS	210,193			
	(1,908)			

La inclusión de (CULT_MED) produce una mejora del ajuste, es decir, aumenta nuestra capacidad predictiva del rendimiento. Sin embargo, el cambio no alcanza la magnitud exigida como mínima para aceptar su permanencia.

* La diferencia entre el valor de verosimilitud de este modelo (=209609,6) es 6,2 puntos inferior al del anterior (=209615,8), con una probabilidad de ocurrencia de 1,04% con 1 gl.

En otros contextos, tal magnitud podría considerarse significativa. Sin embargo, dado que nuestra finalidad es definir el modelo que deberemos utilizar como criterio de "control" en el resto de los análisis, el criterio de la parsimonia parece especialmente relevante. Por ello, preferimos adoptar como modelo final al representado en el Cuadro 3.2 que contiene dos mediciones individuales (NSE y Icul2r2) y la composición socioeconómica de la escuela (NSE_MED).

De acuerdo a estos resultados, podemos concluir que:

a) la medición clásica del nivel socioeconómico familiar, a través de la combinación de la educación de los padres y la disponibilidad de bienes de uso durable, se mostró muy adecuada, avalando tal simplificación en medición;

b) el acceso a los bienes culturales se distingue nítidamente del indicador anterior y ayuda a predecir los niveles de rendimiento. Por lo tanto, es aconsejable medirlo y mantenerlo diferenciado a lo largo del análisis;

c) finalmente, no se puede obviar la inclusión de la medición "contextual", especialmente la referida a la composición social del establecimiento (NSE-MED), que resultó con una capacidad predictiva muy alta, aún después de haber considerado las mediciones individuales.

Este modelo final, conteniendo dos indicadores a nivel del alumno (variables de nivel 1: nivel socioeconómico familiar y acceso a bienes culturales) y uno a nivel de la escuela (variable de nivel 2: composición social de la escuela), será utilizado en todos los análisis subsiguientes para evaluar la eficacia de las diversas variables cuando se "controla" por el origen social del alumno. Por ello, de aquí en más lo llamaremos "**modelo de referencia**".

6. La variación de la equidad educativa institucional. Finalizando este capítulo, nos proponemos responder a la última pregunta planteada al inicio, a saber: ¿todas las escuelas distribuyen el aprendizaje con el mismo nivel de equidad? En este contexto específico, entendemos por **(in)equidad educativa el grado en que la distribución de los aprendizajes acompaña el perfil de las desigualdades en el origen social del alumnos**. Por lo tanto, para responder esta pregunta debemos analizar con mayor detalle el comportamiento de la relación entre (REND) y (NSE).

Anteriormente, pudimos establecer también que el rendimiento promedio de las escuelas, ajustado por (NSE), varía significativamente. Pero en ese análisis, se había supuesto que la intensidad de asociación entre (NSE) y (REND) era similar en todas las escuelas estudiadas. Es decir, se asumía que las diferencias de rendimiento escolar entre los alumnos de alto y bajo nivel socioeconómico son estadísticamente iguales en todas las escuelas. Esto obviamente, puede estar muy lejos de la realidad. Por lo tanto, ahora se abandonará tal supuesto, permitiendo que esa relación varíe de escuela en escuela (aleatorización). Es necesario realizar otro paréntesis metodológico para explicar esta operación.

Paréntesis metodológico 2

En este bloque se exponen algunos conceptos y operaciones implicadas en la elaboración de tal respuesta, y debe considerarse continuación del paréntesis metodológico anterior de este mismo capítulo.

El grado de equidad educativa se define operacionalmente como el valor de la pendiente **b** de la ecuación en (3) (ver Paréntesis metodológico 1); ese parámetro nos indica la intensidad con que (REND) aumenta (disminuye) a medida que (NSE) aumenta (disminuye). En otras palabras, indica la sensibilidad de (REND) respecto a (NSE). A medida que la pendiente sea menos (más) pronunciada, aumentará (disminuirá) el nivel de equidad. Según vimos anteriormente, esa relación de dependencia puede ser representada por una ecuación lineal (o no-lineal) y, si aleatorizamos, habrá tantas ecuaciones como escuelas existan en el sistema. Sin embargo, no todas las ecuaciones se situarán al mismo nivel de rendimiento promedio ni registrarán la misma intensidad o fuerza de relación entre (REND) y (NSE). En otras palabras, existirá una variación de los parámetros **a** y **b** de las diferentes ecuaciones.

Al permitir que las pendientes varíen de escuela en escuela, la ecuación en (3) se re-escribe

$$(REND)_{ij} = a + b_j(NSE)_{ij} + \epsilon_j + e_{ij} \quad (6)$$

donde b_j es la pendiente de la regresión para la j -ésima escuela.

Como hicimos anteriormente con \mathbf{a} , podemos considerar b_j compuesto por dos elementos, o sea, $b_j = \beta_0 + \beta_1 \mu_{1j}$. Resumiendo, \mathbf{a} es el intercepto y \mathbf{b} la pendiente de la línea mediapoblacional (parte fija del modelo), mientras que β_0 y β_1 expresan la cuantificación de la distancia que separa la j -ésima escuela y aquella dos medias poblacionales (parte aleatoria del modelo al nivel 2 o sea, la escuela), y se supone que ambas están distribuidas normalmente, con media igual a cero y varianza a ser estimada. Como ambas varían aleatoriamente en el mismo nivel, pueden estar correlacionadas entre sí y por tanto, es necesario estimar también su covarianza. En consecuencia, el modelo ahora es

$$(REND)_{ij} = (a + b(NSE)_{ij}) + (\beta_0 + \beta_1 \mu_{1j} + e_{ij}) \quad (7)$$

de donde se deberá estimar β_0 , β_1 , σ_{μ}^2 y σ_e^2 . De hecho, al estimar esos parámetros aleatorios se puede conocer la estructura de la variación total no-explicada por el modelo.

O a la inversa, el objetivo es conocer la forma en que se relacionan el nivel socioeconómico del alumno y su rendimiento escolar a nivel de las escuelas. La variación total se expresa como

$$\text{var}(\beta_0 + \beta_1 \mu_{1j}) = \sigma_{\mu}^2 + 2\beta_1 \sigma_{\mu}^2 \mu_{1j} + \beta_1^2 \sigma_{\mu}^2$$

A los fines de una diferenciación visual de la parte aleatoria y fija del modelo, la ecuación en (7), referida a (matema), se puede re-escribir:

$$\begin{aligned} \text{matema}_{ij} &\sim N(\mathbf{X}\mathbf{B}, \Omega) \\ \text{matema}_{ij} &= \beta_{0ij} \text{cons} + \beta_{1j} \text{nse}_{ij} \\ \beta_{0ij} &= \beta_0 + \mu_{0j} + e_{0ij} \\ \beta_{1j} &= \beta_1 + \mu_{1j} \end{aligned} \quad (8)$$

$$\begin{bmatrix} \mu_{0j} \\ \mu_{1j} \end{bmatrix} \sim N(\mathbf{0}, \Omega_{\mu}) : \Omega_{\mu} = \begin{bmatrix} \sigma_{\mu 0}^2 & \\ & \sigma_{\mu 1}^2 \end{bmatrix}$$

$$\begin{bmatrix} e_{0ij} \end{bmatrix} \sim N(\mathbf{0}, \Omega_e) : \Omega_e = \begin{bmatrix} \sigma_{e 0}^2 \end{bmatrix}$$

donde β_0 se denota con β_{0j} , β_1 con β_{1j} y μ_{0j}

con μ_{1j} .

Recordemos que el concepto de "eficacia", presentado al inicio de este capítulo, se sitúa justamente en este campo de variabilidad y está constituido por dos dimensiones, a saber: **la calidad y la**

equidad. El primer aspecto se refiere al puntaje promedio obtenido por una escuela o sección en la prueba de conocimiento, habiéndose previamente controlado el nivel socioeconómico del alumno. Los resultados correspondientes ya fueron oportunamente analizados. Por lo tanto, al enfocar el comportamiento de la equidad estaremos completando el análisis de la eficacia.

Retomemos ahora el "modelo final" identificado en el paso anterior (ver cuadro 3.2). Si lo liberamos del supuesto de equidad (o pendientes) iguales en todas las escuelas, transformamos a ésta en una variable a ser estudiada, respecto a la cual debemos estimar su variabilidad. Además, podemos preguntarnos si existe alguna **interacción** entre rendimiento institucional y equidad.

Al ejecutar este modelo, las estimaciones en la parte fija no se alteraron. Por lo tanto, la atención será dirigida exclusivamente a su parte "aleatoria". Los resultados se presentan en el Cuadro 3.4 y admiten los siguientes comentarios:

3.4 - Parte aleatoria con (NSE)		
CONS	σ_{u0}^2	
	122,096 (5,935)	
NSE	σ_{u10}	σ_{u1}^2
	0,111 (0,057)	0,006 (0,001)

Test de máxima verosimilitud = 209561,8

. En principio, el modelo aleatorio es estadísticamente significativo;

* El valor del test de máxima verosimilitud de este modelo es 209561,8, distante del modelo "referencia" en 54 puntos, diferencia que con 2g.l., tiene una probabilidad de ocurrencia muy inferior al 1 por mil;

. El término (NSE)/(NSE), que denota la variación del grado de asociación entre (REND) y (NSE) a nivel de las escuelas, es estadísticamente significativo; por lo tanto, podemos afirmar que la diferenciación social en la distribución de los rendimientos varía fuertemente entre las escuelas;

existen escuelas más equitativas que otras. Es decir, la distribución del aprendizaje independientemente del origen social del alumno no es similar en todas las escuelas;

. Del signo positivo (+) del término (CONS)/(NSE), que denota la interacción entre el término (NSE)/(NSE) (beta) y el rendimiento promedio de la escuela, podría inferirse que a medida que se aumenta el rendimiento promedio, se observará mayor asociación entre (NSE) y (REND). Sin embargo, el valor de la estimación nos llevó a sospechar de su nivel de significación. Al contrastarlo con el modelo que excluye este término y mantiene sólo a (NSE)/(NSE), nuestras sospechas se vieron confirmadas. Por lo tanto, debemos descartar aquella interpretación.

?? Si se extrae (CONS)/(NSE) y se reprocesa el modelo, se obtiene 209565,2 como valor de máxima verosimilitud, o sea, una diferencia de apenas 3,4 puntos con el modelo que estamos analizando. Esa diferencia tiene una probabilidad asociada de ocurrencia de 0,065196, es decir, muy por arriba de los criterios convencionales; por lo tanto, debemos concluir que no existe covarianza entre ambos términos.

Por lo tanto, el modelo más eficiente (ajustado y parsimonioso) para representar las relaciones entre el factor socioeconómico y el rendimiento es el expuesto en el Cuadro 3.5. En otras palabras, este es el modelo que permite **cuantificar la calidad y la equidad educativas de la forma más ajustada y sintética posible**, de acuerdo a los datos del ONE 1997.

3.5 - Modelo Final: Rendimiento y Nivel			
CONS	NSE	ICUL2R2	NSE_MED
48,611 (0,374)	0,062 (0,005)	0,065 (0,005)	0,308 (0,021)
☺	CONS	NSE	
CONS	122,647 (5,955)		
NSE		0,007 (0,001)	
☺	CONS		
CONS	207,321 (1,915)		

Rendimiento y nivel socioeconómico en diferentes años. Deseamos cerrar este capítulo respondiendo a una legítima curiosidad: ¿(la estimación de) el efecto del origen social del alumno ha sido estable en las diversas mediciones realizadas durante la década? En principio, la respuesta es afirmativa. El efecto sobre ambas áreas, Matemática y Lengua, ha sido el siguiente:

	Año	Matemática	Lengua
	1994	14,6	16,1
	1995	11,9	14,5
	1996	11,5	11,2
	1997	12,5	14,1

Las leves oscilaciones observadas deben adjudicarse al carácter muestral de las estimaciones y a pequeños cambios en la medición de (nse). Pero, en general, estos datos refuerzan la confiabilidad de las estimaciones presentadas a lo largo de este capítulo y en especial, las referidas a la magnitud del efecto que el origen social del alumno tiene sobre su aprendizaje escolar.

CONCLUSIONES

Los resultados obtenidos al analizar los datos producidos por el ONE de 1997 indican que el origen social del alumno tiene un peso importante en la explicación de las diferencias en los rendimientos en Matemática. Se pudo constatar que tal incidencia se expresa no solamente a través de las desigualdades de aprendizaje entre los alumnos, sino también y principalmente, a través de las diferencias de rendimiento promedio entre las escuelas. La hipótesis de que el sistema educativo argentino se encuentra socialmente segmentado ha sido frecuentemente planteada. Los datos del ONE permiten no sólo corroborar esta idea, sino también dimensionarla con mayor precisión.

Pero, por otro lado, fue posible establecer que la variación de los rendimientos promedios de las escuelas **continúa siendo significativa**, aún después de haber controlado el factor socioeconómico. Es decir, los datos no avalan una visión determinista para la explicación de las desigualdades en el aprendizaje. Existe una variación significativa de la **calidad educativa institucional entre las escuelas con la misma composición social**. Otro tipo de factores debe ser incorporado al análisis de las posibles causas de las desigualdades en el aprendizaje de los alumnos y en la eficacia de las escuelas.

De la misma forma, se pudo observar que **existen escuelas más equitativas que otras**. Los niveles de inequidad educativa institucional varían significativamente entre las escuelas. Algunas de ellas consiguen, quizás a través de la implementación de prácticas (institucionales o pedagógicas), que la distribución de los saberes escolares esté menos atada al condicionamiento social del alumnos.

Desde el punto de vista instrumental, ha sido probada la eficiencia del índice de nivel socioeconómico construido con informaciones provenientes del cuestionario del alumno. Además, el comportamiento de los datos indica la conveniencia del uso simultáneo del índice a nivel individual y contextual (composición social de la escuela). La medición del acceso de los bienes culturales se mostró parcialmente independiente del índice de nivel socioeconómico, avalando su uso para mejorar la bondad de las predicciones del rendimiento a partir del contexto socioeconómico.

Capítulo 4

RENDIMIENTO Y CARACTERÍSTICAS PERSONALES DEL ALUMNO

Hasta aquí, el análisis ha considerado características de la familia del alumno. Abordaremos ahora las relaciones entre algunas características propias de alumno y el rendimiento obtenido en Matemática. Disponemos de informaciones sobre el género, la edad, repitencia, el cambio de escuela, el tiempo y la distancia a la escuela. No podemos afirmar, sin embargo, que todas estas variables se definen a nivel del alumno "per se". Sin dudas, las dos primeras pueden ubicarse en esa categoría, pero la repitencia, el cambio y la lejanía a la escuela son características que se sitúan a nivel de la relación alumno-escuela.

Objetivos. Nos proponemos determinar cuáles de esas características del alumno mantienen una asociación estadísticamente significativa con el rendimiento. Como fuera advertido oportunamente, esa tarea se realizará considerando también a las variables referidas al origen socioeconómico del alumno y a la composición social de la escuela, es decir, en relación al modelo de "referencia" definido en el capítulo anterior (ver Capítulo 3, Cuadros 3.2 y 3.5). Por lo tanto, se intentará responder a las siguientes preguntas específicas:

1. El sexo, la repitencia escolar, el cambio de escuela y la distancia (espacio-temporal) a la escuela ¿contribuyen a explicar las diferencias en el rendimiento escolar?

2. La posible asociación de esas variables con el rendimiento, ¿se mantiene aún después de haber **controlado el nivel socioeconómico del alumno y la composición social de la escuela?**
3. El efecto de cada uno de los factores hipotetizados ¿se mantiene aún cuando todos ellos actúen conjuntamente?
4. Si tal efecto existe, ¿su intensidad varía de escuela en escuela?
5. ¿Existe alguna interacción entre tales factores?

Se busca pues, precisar cuál/es de los factores individuales del alumno mantiene/n una asociación significativa con el rendimiento escolar, mas allá del efecto que tienen el nivel socioeconómico del alumno, el contexto socioeconómico escolar y los otros factores individuales considerados.

Como se verá posteriormente, la primera pregunta se refiere a la determinación del efecto principal de cada variable sobre el rendimiento escolar, teniendo en cuenta la incidencia del nivel socioeconómico del alumno. La segunda en cambio, exige conocer la alteraciones en la varianza "entre-escuela" cuando se considera cada factor. La tercera pregunta requiere especificar el modelo que mejor ajusta a los datos cuando se deja que las relaciones entre las variables fluctúen libremente entre las escuelas y entre los alumnos. Por último, se deberá evaluar si la asociación de cada factor escolar con el rendimiento persiste aún cuando se tomen en cuenta conjuntamente todos los otros factores individuales del alumno.

Las variables¹¹. Las variables que consideraremos en este capítulo, además del rendimiento y las que componen el "modelo de referencia" (ver Capítulo 3, Cuadro 3.2), son las siguientes:

- (REPITIEN)** En el cuestionario correspondiente, al alumno se le preguntó si alguna vez había repetido de año. En base a esa información se pudo clasificar al alumno como repitiente o no-repitiente. Para el análisis, es necesario crear una variable "muda" con referencia a alguna de las dos categorías. Elegimos la categoría repitiente (ver Paréntesis metodológico 3 más adelante);
- (EDAD)** La edad del alumno (variable intervalar);
- (FEMENINO)** La variable "muda" es la categoría femenino; la "base" es masculino;
- (CAMBIARO)** Al alumno se preguntó cuántas veces había cambiado de escuela en los dos últimos años. La variable "muda" es haber cambiado alguna vez; la "base" es no haber cambiado de escuela;
- (TIEM*DIS)** El alumno fue interrogado acerca de cuánto tiempo demora en llegar a la escuela desde su casa y a qué distancia vive de la escuela. A los fines de la simplicidad, con estas dos mediciones se construyeron dos índices, uno sumativo y el otro multiplicativo. Las pruebas preliminares con ambos demostraron que el segundo era más eficaz (más predictivo) respecto al rendimiento, por lo cual decidimos usar éste y desechar el otro.

¹¹ Ver Cuestionario del alumno en Anexo .

Estrategia metodológica. Tal cual venimos haciéndolo, los datos serán analizados con la técnica estadística de "niveles múltiples". Para responder a las preguntas planteadas, seguiremos el siguiente camino:

1. Estimación del modelo "vacío";
2. Análisis de la asociación entre el rendimiento y el sexo, la edad, la repitencia, el cambio de escuela y la lejanía espacio-temporal a la escuela;
3. estimación del modelo de "referencia";
4. cada uno de los factores se evalúa con el modelo "referencia";
5. cada factor se evalúa con el modelo de "referencia" pero incluyendo también a los otros factores;
6. finalmente, análisis de aleatorización de los factores a nivel de escuela (nivel 2);

Muestra y test de significación. El archivo de datos analizados es de 912 escuelas con 24.849 alumnos. Para evaluar la asociación entre factor y rendimiento continuaremos usando la razón de máxima verosimilitud (ver Capítulo 1), con probabilidad igual o menor al 1 por mil.

Resultados. En el cuadro 4.1 se presentan los resultados obtenidos en los dos primeros pasos, con los parámetros necesarios para la comprensión del análisis (la estimación, el valor del test de máxima verosimilitud, la razón respecto al modelo "vacío" y el nivel de significación estadística). Antes de comentar esos resultados, haremos un paréntesis metodológico referido al método de análisis de las variables "dummy". Esta nota metodológica puede considerarse continuación de las anteriores (ver Capítulo 3), donde se trató el análisis de las variables intervalares.

Paréntesis metodológico 3

Definición y análisis de variables "dummy".

A todas las variables analizadas en los capítulos anteriores las hemos considerado intervalares. Sin embargo, a varios de los factores que analizaremos a lo largo de este trabajo los debemos tratar como variables nominales o categoriales. Para poder incluir este tipo de variables en el análisis, debemos definirlos como "clasificaciones fijas", usando el método de variables "mudas". En primer lugar, respecto a una variable cualquiera (ej. repitencia), cada observación (en este caso, cada alumno) debe estar asignada en una y sólo en una categoría (ej. alumno no-repitente o alumno repitente). A seguir, habrá que decidir cuál será la "categoría base" (digamos, "no-repitente") y como consecuencia, las categorías restantes serán variables "muda" (ej. repitente).

Para explicar cómo se analiza este tipo de variable, usaremos como ejemplo a la variable (REPITIEN). Hacemos que, para cualquier observación (alumno), el valor de la variable "muda" correspondiente sea 1 si pertenece a la categoría de repitente y 0 si no pertenece ($z_{ij} = 1$, si es repitente; $z_{ij} = 0$, si no lo es). Generalizando, tendremos tantas variables "mudas" como el número de categorías, menos una (correspondiente a la "categoría base").

Veamos cómo se representa el efecto de una variable de "clasificación fija" sobre el rendimiento escolar. Volvamos a la ecuación (5) en el Capítulo 3. Re-escribimos esa fórmula de tal forma que exprese la relación entre rendimiento y la situación de repitencia:

$$\begin{aligned} \text{matema}_{ij} &\sim N(\chi B, \Omega) \\ \text{matema}_{ij} &= \beta_{0ij} \text{cons} + \beta_1 \text{repitien}_{ij} \\ \beta_{0ij} &= \beta_0 + u_{0j} + e_{0ij} \\ \begin{bmatrix} u_{0j} \end{bmatrix} &\sim N(0, \Omega_u) : \Omega_u = \begin{bmatrix} \sigma_u^2 \end{bmatrix} & (9) \\ \begin{bmatrix} e_{0ij} \end{bmatrix} &\sim N(0, \Omega_e) : \Omega_e = \begin{bmatrix} \sigma_e^2 \end{bmatrix} \end{aligned}$$

donde $(\text{matema})_{ij}$ es el rendimiento en matemática del alumno i en la escuela j ; β_0 es la estimación del rendimiento promedio de los alumnos no-repitientes ("base"); β_1 es la diferencia esperada entre el rendimiento de los alumnos repitientes y no-repitientes, es decir, una estimación de la ventaja que tienen los alumnos no-repitientes sobre los repitientes; y $(\text{repitien})_{ij}$ es el valor de la variable "muda" del alumno i en la escuela j .

El procesamiento de los datos en base a esta metodología, arrojó los siguientes resultados para la variable de repitencia (ver Cuadro 4.1):

$$(\text{REND})_{ij} = 52,099 + (-7,531 z_{ij})$$

Es decir, la diferencia esperada entre los rendimientos de alumnos repitientes y no-repitientes es de 7,531 puntos a favor de estos últimos y el rendimiento promedio de los no-repitientes (variable "base") es igual a 52,099. El test de máxima verosimilitud de este modelo es 206444,40, distante en 871,3 puntos respecto al del modelo "vacío" (=207315,700), diferencia que, con 1 gl., tiene una probabilidad de ocurrencia mucho menor al 1 por mil.

Final del paréntesis metodológico

1. El modelo "vacío". Las estimaciones no difieren de las obtenidas en los modelos "vacíos" anteriores. Debemos prestar atención solamente al valor del test de verosimilitud (207315,70) propio de este modelo, ya que es la estimación de referencia para evaluar la fuerza de la relación de cada uno de los indicadores con el rendimiento.

2. Las variables individuales. Exceptuando (FEMENINO), las relaciones de todas las variables con el rendimiento arrojaron valores en el test de máxima verosimilitud significativamente distantes del correspondiente al modelo "vacío". El signo de la estimación de todas esas relaciones es negativo (-). Dada la definición de las variables, se puede inferir que:

~~el~~ el puntaje esperado de los alumnos repitientes es significativamente menor (-7,5 puntos porcentuales) al de los no repitientes;

~~la~~ medida que aumenta la edad del alumno (sobre edad), disminuye su rendimiento;

~~el~~ el puntaje de los alumnos que cambiaron de escuela es significativamente menor (-2,95 puntos porcentuales) al de aquellos que no lo hicieron;

~~cuanto~~ cuanto más distante de su escuela vive el alumno, menor es su rendimiento;

finalmente y sorprendentemente, existe un indicio de que las alumnas obtienen más bajo rendimiento (-0,58 puntos) que los alumnos.

Cuadro 4.1 - Parámetros estimados para indicadores de características personales del alumno.

Parámetro	Estimación militud	Verosi	Razón Sig.	Nivel
1. Modelo "vacío" - 207315,700				
2. Variables individuales				
(FEMENINO)	-0,582	207307,00	8,7	**
(REPITIEN)	-7,531	206444,40	871,3	***
(EDAD)	-2,451	206952,00	363,7	***
(CAMBIARO)	-2,955	207176,40	139,3	***
(TIEM*DIS)	-0,452	207256,30	59,4	***

Nota: (**): prob. 1%; (***): prob. 1 por mil;

3. Modelo de referencia. El tercer paso consiste en definir el modelo que hemos denominado de "referencia" (ver Cuadro 3.2 en el Capítulo 3). En el Cuadro 4.2 se presentan los resultados de esa operación. Al igual que en el caso del modelo "vacío", las estimaciones obtenidas fueron coincidentes con las exhibidas en los capítulos precedentes. Para nuestros fines, debemos retener el valor del test de máxima verosimilitud (=206416,00; Cuadro 4.2), ya que sirve como valor de referencia para evaluar la fuerza de la relación de cada uno de los indicadores con el rendimiento, una vez controlada la acción de las variables socioeconómicas. El rendimiento promedio estimado en este modelo es 43,487%.

4. Características personales y nivel socioeconómico. A continuación se evalúa individualmente cada uno de los factores con el modelo de "referencia" definido anteriormente. Las estimaciones resultantes se observan en la parte 4 del Cuadro 4.2.

. Con excepción de (FEMENINO), todas las razones de máxima verosimilitud respecto al modelo de "referencia" son altamente significativas (prob. 1 por mil). Por lo tanto, la situación de repitiente, la edad, el cambio y lejanía de escuela tienen incidencia sobre el rendimiento escolar, inclusive después de extraer la proporción de la variación explicada por las variables socioeconómicas. En otras palabras, estos factores agregan un efecto "neto", más allá del que tienen las variables del modelo de "referencia";

* Obsérvese que el rendimiento promedio estimado para la categoría "base" de cada variable es superior al estimado en el modelo "referencia" para toda la población (43,487%). Por ejemplo, el rendimiento promedio estimado para los no repitientes (46,061% es superior en 2,584 puntos porcentuales del estimado para toda la población.

. Sin embargo, es importante observar que una parte de la distancia entre los rendimientos debida a algunas de estas variables podían estar expresando, en realidad, la operación de los factores socioeconómicos. Así, observamos por ejemplo, que la distancia entre los rendimientos esperados de repitientes y no-repitientes disminuye de 7,531 (Cuadro 4.1) a 6,363; el peso de la incidencia de la lejanía a la escuela desciende de 0,452 a 0,227;

. Finalmente, la distancia de rendimiento entre las categorías masculino/femenino no es tan nítida como la verificada en las otras variables estudiadas. Las alumnas tiene un rendimiento esperado apenas 0,431 puntos porcentuales inferior al de los varones. Sin embargo, decidimos explorar ese indicio de asociación y mantenerlo como válido en el análisis subsecuente.

* La diferencia entre el valor de verosimilitud entre este modelo (=206411,2) con el de "referencia" es 4,8 puntos que, con 1gl., tiene una probabilidad de ocurrencia de 0,02846, o sea, muy lejano del criterio más estricto del 1 por mil.

Cuadro 4.2 - Parámetros estimados para indicadores de características personales del alumno en el modelo de "referencia".

Parámetros	Promedio estimado (β_0)	Estimación diferencia (β_1)	Verosi militud	Razón	Nivel Sig.
3. Modelo "referencia"					
	43,487	206416,0	-	-	
4. Variables individuales					
(REPITIEN)	46,061	-6,363	205803,20	612,0	***
(CAMBIARO)	44,221	-2,949	206273,30	142,7	***
(FEMENINO)	43,642	-0,431	206411,20	4,8	*
(EDAD)	51,098	-1,876	206204,30	211,7	***
(TIEM*DIS)	44,274	-0,227	206377,60	38,4	***

(*): prob.?5%; (**): prob.?1 por mil;

Nota: Los rendimientos promedios estimados se refieren a la categoría "base: no-repitente", no cambiaron de escuela y sexo masculino.

5. El conjunto de los factores. Este paso consiste en incluir los factores simultáneamente en el modelo de referencia. De esta forma, se podrá evaluar si el aporte individual de cada factor se

mantiene aún después de haber considerado a los restantes. En el Cuadro 4.3 se presentan los principales resultados de esta operación.

. en primer lugar, experimentamos la acción conjunta de la edad y la situación de repitencia. Podemos observar que la estimación de la edad ha descendido abruptamente, quedando muy próxima a su error standard. Concluimos que la incidencia de esta variable sobre el rendimiento se superpone con la sostenida por la repitencia y por lo tanto, no adiciona nada a la predicción del rendimiento;

* Esta falta de significación se comprueba porque el valor del test de máxima verosimilitud de este modelo (=205802,8) se distancia en sólo 0,4 puntos del correspondiente al modelo que incluye a (REPITIEN) sólo (=205803,2; ver Cuadro 4.2). Adviértase que el nivel de significación de un modelo con 2 factores se basa en la comparación de su eficacia predictiva con la del factor más eficaz de los dos.

. A continuación, incluimos el cambio de escuela (CAMBIARO) junto con (REPITIEN). Observemos que las estimaciones de la distancia entre alumnos repitientes y no repitientes (-6,115 puntos) y de los que cambiaron y no cambiaron de escuela (-2,393), permanecen con valores próximos a los del modelo anterior (ver Cuadro 4.2); por lo tanto, el aporte de cada una de estas variables como predictora del rendimiento es significativo e independiente del de la otra.

* De hecho, el valor que arroja el test de máxima verosimilitud (=205709,3) dista del modelo anterior en 93,5 puntos, con una probabilidad de ocurrencia muy inferior al 1 por mil. Consistentemente, resultados no presentados aquí permitieron constatar que las estimaciones de ambas variables son muy distantes del error standard correspondiente

En este modelo, el rendimiento promedio estimado (=46,553) se refiere a los alumnos que reúnen las dos condiciones implicadas, es decir, alumnos "no-repitientes" y que "no cambiaron" de escuela en los dos últimos años. Cuando se consideraba solamente la dicotomía repitiente/no-repitiente, el promedio estimado para los no-repitientes fue 46,061% y por tanto, inferior al de los alumnos que adicionalmente, cumplen la condición de no haber cambiado de escuela.

. Ahora procedemos a incluir en el análisis la diferencia de rendimiento entre alumnos y alumnas. Aparentemente, ahora la distancia entre ambas categorías se ha profundizado: las mujeres tienen un rendimiento esperado 0,806 puntos porcentuales inferior al de los varones y esa diferencia es estadísticamente significativa.

* La diferencia entre el test de máxima verosimilitud de este modelo y el anterior es de 16,6 puntos que, con 1 gl., tiene asociada una probabilidad inferior al 1 por mil.

Ese cambio se produce porque ahora estamos comparando a los alumnos no sólo en base a su sexo, sino también a su situación de repitencia y cambio de escuela.

. Finalmente, inspeccionamos qué sucede con la asociación entre el rendimiento y la variable referida a la lejanía geográfica del alumno respecto a la escuela. Aún cuando la sensibilidad del rendimiento respecto a esa variable disminuye, la magnitud de la estimación respecto al error standard continúa indicando una asociación muy significativa entre ambas variables.

* El valor del test de máxima verosimilitud (=205666,5) es 26,2 puntos menor que el del modelo anterior, con una probabilidad de ocurrencia muy inferior al 1 por mil cuando se tiene 1 gl.

* Es interesante observar también que ahora la línea que representa la relación entre rendimiento y distancia a la escuela, definida por los parámetros β_0 (47,307) y β_1 (-0,182), corresponde a los alumnos en las categorías "base", es decir, no-repitientes, sin cambio de escuela y masculinos. Esos valores reflejan no sólo un promedio esperado mayor, sino también, una relación rendimiento/distancia menor (para la población, el parámetro β_1 fue estimado en 0,227; ver cuadro 4.2)

Cuadro 4.3 - Parámetros estimados para sub-conjuntos de características personales del alumno en el modelo de "referencia".

Parámetros	Promedio	Estimación	Verosi	Razón	Nivel
estimado	diferencia	militud	Sig.		
(β_0)	(β_1)				
(EDAD)		0,093			
(REPITIEN)	45,589	-6,466	205802,8		ns
(REPITIEN)		-6,115			
(CAMBIARO)	46,437	-2,393	205709,3	93,5	***
(REPITIEN)		-6,182			
(CAMBIARO)		-2,396			
(FEMENINO)	46,736	-0,806	205692,7	16,6	***
(REPITIEN)		-6,150			
(CAMBIARO)		-2,344			
(FEMENINO)		-0,775			
(TIEM*DIS)	47,307	-0,182	205666,5	26,2	***

(***): prob. ?1 por mil; (ns): no significativo.

De esta forma, queda definida la Parte Fija del modelo que mejor representa las relaciones entre algunas características personales del alumno y el rendimiento. Cada una de ellas disminuye una proporción estadísticamente significativa de la variación no-explicada del rendimiento, aún cuando se controle no sólo a las otras características personales incluidas, sino también a las variables que constituyen el modelo de "referencia". Por esto último, podemos afirmar que la diferencia de rendimiento estimada entre, digamos, alumnos repitientes y no repitientes, es válida bajo el supuesto de igualdad en el nivel socioeconómico y de acceso a bienes culturales del alumno, y en la composición social de la escuela. Las estimaciones de este modelo (parte fija y aleatoria se presentan en el Cuadro 4.4.

4.4 - Modelo final - Matemática							
CONS	NSE	BIE_CULT	NSE_MED	REPITIEN	CAMBIARO	FEMENINO	TIEM*DIS
47,471	0,048	0,057	0,291	- 6,158	- 2,334	- 0,758	- 0,184
(0,665)	(0,005)	(0,005)	(0,022)	(0,257)	(0,245)	(0,193)	(0,036)

CONS	119,129	
	(5,959)	
CONS	207,855	
	(1,900)	

6. Aleatorización a nivel de escuela. Hasta aquí habíamos supuesto que la intensidad de asociación entre cada factor y el rendimiento era similar en todas las escuelas estudiadas. Pero ahora, nos interesa saber si tales efectos varían de escuela en escuela. Por ejemplo, ya no supondremos que las diferencias de rendimiento entre alumnos repitientes y no repitientes son estadísticamente iguales en todas las escuelas. Por el contrario, abandonaremos ese supuesto, permitiendo que las relaciones varíen de escuela en escuela (aleatorización).

* **Nota metodológica.** En el Cap. 3 introdujimos el análisis de aleatoriedad de la variable intervalar (NSE) a nivel de escuela (nivel 2) (ver ecuaciones (7) y (8) en el Capítulo 3). Ahora desarrollaremos ese análisis con una variable nominal. Para estimar los parámetros aleatorios a nivel de escuela (nivel 2), introducimos la varianza del término a evaluar (ej. (REPITIEN) en el modelo en (9) de este Capítulo). Veamos cómo se expresa formalmente esta operación:

$$\text{matema}_{ij} \sim N(XB, \Omega)$$

$$\text{matema}_{ij} = \beta_{0ij} \text{cons} + \beta_{1j} \text{repitien}_{ij}$$

$$\beta_{0ij} = \beta_0 + u_{0j} + e_{0ij}$$

$$\beta_{1j} = \beta_1 + u_{1j}$$

$$\begin{bmatrix} u_{0j} \\ u_{1j} \end{bmatrix} \sim N(0, \Omega_u) : \Omega_u = \begin{bmatrix} \sigma_{u0}^2 & 0 \\ 0 & \sigma_{u1}^2 \end{bmatrix}$$

$$\begin{bmatrix} e_{0ij} \end{bmatrix} \sim N(0, \Omega_e) : \Omega_e = \begin{bmatrix} \sigma_{e0}^2 \end{bmatrix}$$

El nuevo término σ_{u1}^2 representa la varianza "entre-escuela" de la distancia repitiente/no-repitiente en el rendimiento

En el análisis subsiguiente evaluaremos la aleatorización de cada variable en relación al modelo definido en el paso anterior (Cuadro 4.3), adicionando la aleatorización de (NSE) a nivel de escuela, dada su relevancia estadística demostrada en el Cap. 3 (Cuadro 3.5). Es decir, se trata del modelo que especifica tanto la influencia **fija** que tienen las variables socioeconómicas y personales del alumno sobre el rendimiento, cuanto la variación "entre-escuela" del efecto de (NSE). En el Cuadro 4.5 se muestran las estimaciones de su Parte aleatoria. Es importante que retengamos el valor del test de máxima verosimilitud (=205637,4).

Las estimaciones de la Parte fija que arrojó este modelo fueron similares a las correspondientes al modelo final del paso anterior (Cuadro 4.4), por lo cual nos eximimos de su presentación y análisis. Por otro lado, las estimaciones de la parte aleatoria son consistentes con las expuestas en el Cuadro 3.5 del Cap. 3 y por lo tanto, se mantienen las conclusiones extraídas oportunamente a este respecto.

4.5 - Aleatorización (NSE) Escuela .		
	CONS	NSE
CONS	118,245 (5,976)	
NSE		0,005 (0,001)
Test de máxima verosimilitud: 205637,4		

A partir de aquí, la exposición está organizada en tres momentos: en el primero, comentaremos los resultados de la variación del efecto de cada variable; en seguida, se investiga la posible interacción de cada variable con el efecto de (NSE); finalmente, estudiamos las posibles interacciones entre las propias variables.

6.1. La variación de los efectos de las variables "dummy"

(REPITIEN) En el Cuadro 4.6 se presenta la estimación de la variación "entre-escuela" del efecto de la variable repitencia. Los resultados indican que la distancia entre repitientes y no-repitientes no es significativamente diferente entre las escuelas. En principio, no habría diferencias entre las instituciones en cuanto a la capacidad de promover mayor equidad en la distribución de aprendizajes, respecto a la "historia escolar" del alumno.

4.6 - Aleatorización (NSE) y (REPITIEN) Escuela

El valor de verosimilitud de este modelo (=205634,9) se distancia apenas en 2,5 puntos respecto al modelo adoptado como "referencia" en el análisis de la aleatorización, valor cuya probabilidad de ocurrencia es superior al 5% cuando se tiene 1 gl.

☺	CONS	NSE	REPITIEN
CONS	117,981 (5,974)		
NSE		0,005 (0,001)	
REPITIEN			3,547 (2,411)

(CAMBIARO) Seguimos el mismo procedimiento aplicado anteriormente, pero ahora con la variable que indica el cambio de escuela. Los resultados (Cuadro 4.7) no nos permiten inferir conclusiones tan nítidas como en el caso anterior. Si bien la fuerza de la asociación entre ambas variables (rendimiento y cambio) no alcanza el nivel establecido como requisito mínimo, no podemos dejar de notar que existen indicios de que la distancia entre alumnos que cambiaron y no cambiaron de escuela varía entre las escuelas.

El valor de verosimilitud de este modelo (=205632,3) se distancia en 5.1 puntos del correspondiente al modelo adoptado como "referencia". Para 1 gl., esta distancia tiene una probabilidad de ocurrencia de 0,023926, es decir, superior al 1%, pero inferior al 5%.

4.7 - Aleatorización (NSE) y (CAMBIARO) - Nivel escuela			
☺	CONS	NSE	CAMBIARO
CONS	118,182 (5,972)		
NSE		0,005 (0,001)	
CAMBIARO			5,093 (2,481)

(FEMENINO). Finalmente, incluimos la varianza del término referido al sexo del alumno, indicado como (FEMENINO). Ahora, pretendemos saber si las distancias de los rendimientos de varones y mujeres es similar en todas las escuelas o por el contrario, varían. Observemos los resultados obtenidos (Cuadro 4.8). Aparentemente, con esta variable no hay dudas: la distancia entre alumnos y alumnas respecto al rendimiento varía significativamente entre las escuelas. Al contrario de lo ocurrido con las dos variables anteriores, habría diferencias entre las instituciones en cuanto a su capacidad de promover mayor equidad en la distribución de aprendizajes respecto al género.

El valor de verosimilitud de este modelo (=205527,1) se distancia 110,3 puntos respecto al modelo adoptado como "referencia", valor cuya probabilidad de ocurrencia es muy inferior al 1 por mil cuando se tiene 1 gl.

4.8 Aleatorización (NSE) y (FEMENINO) - Nivel Escuela			
☺	CONS	NSE	FEMENINO
CONS	117,578 (6,083)		
NSE		0,004 (0,001)	
FEMENINO			18,192 (2,305)

Otra pregunta que es posible responder, se refiere a la posible asociación entre el rendimiento promedio de la escuela, nominado como (CONS), y la fuerza de la relación rendimiento/género. Deseamos saber si la distancia entre los rendimientos de alumnos y alumnas es más o menos estrecha, según que el rendimiento promedio de la institución sea más (o menos) alto.

Esta operación consiste en introducir el término de la covarianza entre el efecto de (FEMENINO) y el rendimiento promedio de la escuela. Su expresión formal se muestra a la derecha. Para una comparación directa con la que expresa la inclusión del término de varianza (ecuación en (10)), el ejemplo se refiere a (REPITIEN), como único término, tanto en la parte fija como en la aleatoria. El único término nuevo es γ_{10} , denotando la covarianza entre (cons) y (repitien).

$$\begin{aligned} \text{matema}_{ij} &\sim N(XB, \Omega) \\ \text{matema}_{ij} &= \beta_{0ij} \text{cons} + \beta_{1ij} \text{repitien}_{ij} \\ \beta_{0ij} &= \beta_0 + u_{0ij} + e_{0ij} \\ \beta_{1ij} &= \beta_1 + u_{1ij} \\ \begin{bmatrix} u_{0ij} \\ u_{1ij} \end{bmatrix} &\sim N(0, \Omega_u) : \Omega_u = \begin{bmatrix} \sigma_{u0}^2 & \\ & \sigma_{u1}^2 \end{bmatrix} \\ \begin{bmatrix} e_{0ij} \end{bmatrix} &\sim N(0, \Omega_e) : \Omega_e = \begin{bmatrix} \sigma_{e0}^2 \end{bmatrix} \end{aligned}$$

Los resultados obtenidos se pueden observar en el Cuadro 4.9. El nuevo parámetro se designa como (CONS)/(FEMENINO). El comportamiento de los datos indica que existe una correlación significativa entre el nivel de rendimiento y el efecto del género, es decir, la intensidad del efecto del sexo sobre el rendimiento varía en función del promedio de rendimiento de la escuela a la que se asiste. Por la definición de la variable (base=varones) y por el signo negativo de la estimación (-12.793), se infiere que los alumnos obtienen rendimientos más distantes (mejores) de los conseguidos por las alumnas a medida que el rendimiento promedio de la escuela baja. Entonces, los datos de 1997 sugieren que las escuelas que tienen un rendimiento promedio más alto logran distribuir los aprendizajes entre alumnos y alumnas de forma más equitativa.

El valor de verosimilitud de este modelo es igual a (-205507,6), con una diferencia de 19,5 puntos respecto al modelo anterior (Cuadro 4.8), valor cuya probabilidad de ocurrencia es inferior al 1 por mil cuando se tiene 1 gl. Dada su significación estadística, mantendremos este término de covarianza en lo que resta del análisis.

4.9 - Aleatorización de la correlación entre (femenino) y (cons)

	CONS	NSE	FEMENINO
CONS	125,596 (6,729)		
NSE		0,005 (0,001)	
FEMENINO	- 12,793 (3,104)		21,141 (2,549)

Por lo tanto, la diferenciación en la distribución de los aprendizajes según el sexo depende del nivel del rendimiento promedio de la sección o escuela. Saber que una escuela consigue un rendimiento por arriba del esperado ya nos está indicando que la distancia entre los rendimientos de alumnos y alumnas será muy probablemente menor que la del conjunto de todas las escuelas, es decir, esas escuelas son más equitativas que las restantes respecto a ese criterio distributivo (sexo).

6.2. La correlación entre el efecto de las variables "dummy" y el efecto del nivel socioeconómico del alumno.

Corresponde ahora investigar si existe alguna asociación (covarianza) entre el efecto de las categorías que son objeto de este capítulo (repetencia, género y cambio de escuela) y el efecto del nivel socioeconómico del alumno (NSE) sobre el rendimiento.

Comenzamos introduciendo en el modelo anterior (Cuadro 4.9), el término que indica la covarianza entre el efecto de (FEMENINO) con el de (NSE). Con esto, pretendemos saber si la asociación de (NSE) con el rendimiento varía conjuntamente con el grado de diferenciación de los rendimientos de varones y mujeres, y en qué sentido. Los resultados obtenidos se presentan en el Cuadro 4.10 e indican que se debe rechazar la hipótesis de asociación.

El valor de verosimilitud de este modelo es igual a 205506,5 y por tanto, con una diferencia de apenas 1.1 puntos; con 1 gl., esa diferencia tiene una probabilidad asociada superior al 5%.

4.10 - Aleatorización de la covarianza entre (nse) y (femenino)

	CONS	NSE	FEMENINO
CONS	125,556 (6,728)		
NSE		0,005 (0,001)	
FEMENINO	- 13,038 (3,106)	0,038 (0,034)	21,396 (2,566)

Por lo tanto, no se puede aceptar la idea de que en las escuelas donde el rendimiento de los alumnos es más dependiente de su nivel socioeconómico sean al mismo tiempo, en las que la diferenciación varones/mujeres es más pronunciada.

El mismo procedimiento fue aplicado a las otras dos características restantes (repetencia y cambio de escuela) y los resultados fueron igualmente negativos. Por lo tanto, la conclusión anterior, referida al género, puede ser extendida a esas dos variables.

3. La correlación entre el efecto de las variables.

Finalmente, nos abocamos a evaluar si existe variación concomitante entre el efecto de las características personales del alumno que estamos analizando. El procedimiento seguido es el mismo que el del punto anterior, sólo que substituyendo (NSE) por alguna de las variables "dummy". Ninguno de los resultados obtenidos fue estadísticamente significativo, es decir, no pudimos constatar la existencia de correlación entre los efectos de las variables estudiadas del alumno. Esto significa que, por ejemplo, conocer la distancia entre los rendimientos de repitentes y

no-repitientes de una escuela no es una información útil para predecir la magnitud de la distancia entre los rendimientos de varones y mujeres. Lo mismo acontece con el cambio de escuela. Es decir, el grado de equidad que logra una escuela respecto de algunas de las variables individuales del alumno no indica el nivel de equidad logrado en las restantes.

De acuerdo a estos resultados, el modelo que mejor predice el rendimiento del alumno a partir de las características personales estudiadas, en su parte aleatoria contiene la estimación de la variación del efecto del nivel socioeconómico del alumno, de la distancia en el rendimiento entre varones y mujeres y la correlación entre el nivel de rendimiento promedio de la escuela y el efecto del género.

CONCLUSIONES y estabilidad temporal de los resultados.

¿Qué tan estables han sido los resultados de las evaluaciones de calidad educativa en los aspectos tratados en este capítulo? En general, ¿confirman las inferencias establecidas en este capítulo? En general se puede responder que sí, pero existen algunos matices.

1. Respecto de la repitencia no caben dudas: siempre los alumnos repitientes alcanzaron peores rendimientos que los no repitientes. La diferencia de rendimiento (en puntos porcentuales), controlando el efecto de las variables socioeconómicas, fue de 6,196 en 1994; de 4,857 en 1995 y de 5,872 en 1996, a favor de los alumnos no-repitientes, estimaciones muy próximas a la obtenida en 1997 (6,363).

Si bien existe plena consistencia en este aspecto general, no sucede así con la evaluación de la variabilidad "entre-escuela" del efecto de esta característica. En aquellos tres años se detectó una variación significativa de la distancia de rendimiento entre repitientes y no-repitientes a nivel de escuela. En algunas escuelas esa distancia es menor que en otras, es decir, algunas escuelas distribuyen los aprendizajes (entre alumnos repitientes y no-repitientes) de forma más **equitativa**. Dado que esta conclusión está apoyada por los resultados de tres evaluaciones, **parece prudente dejar entre paréntesis la conclusión en contrario extraída de los datos de 1997.**

2. Respecto al cambio de escuela, los datos también son consistentes en cuanto al efecto negativo sobre el rendimiento: los alumnos que cambiaron de escuelas consiguieron rendimiento menor del obtenido por quienes no lo hicieron, en 2,719 (1995), 3,053 (1996) y 2,949 (1997) puntos porcentuales. En esos tres años, el grado de incidencia de esta variable se mostró oscilante entre las escuelas. En ese sentido, las inferencias establecidas anteriormente son bastante sólidas: las escuelas difieren significativamente respecto del efecto del cambio de escuela sobre el rendimiento del alumno.

3. Finalmente, las conclusiones sobre la diferencia de rendimiento según el género parecen las menos terminantes. Mientras que los datos de 1994 y 1996 no mostraban diferencia alguna, en 1995 la diferencia era de .9644. En 1997, observamos una diferencia de 0,431, con probabilidad de 4,8%, es decir, muy cercana al criterio más laxo. Por lo tanto, parece más prudente inclinarse por rechazar provisoriamente la hipótesis de diferencias significativas entre ambos géneros, en el aprendizaje de Matemática.

Respecto de la variación entre-escuela de la distancia masculino/femenino en el rendimiento, los datos parecen ser más terminante: en todos los años pasados esa distancia ha variado significativamente, es decir, existen escuelas que agravan la desigualdad de rendimiento entre género cuando se las compara con otras donde el aprendizaje según género es más equitativa.

Capítulo 5

ACTITUDES, REPRESENTACIONES Y RENDIMIENTO DEL ALUMNO

En este capítulo abordaremos el análisis de las relaciones entre algunas actitudes y representaciones del alumno y su rendimiento en la prueba de Matemática. Más específicamente, nos interesa saber si existe alguna asociación entre el rendimiento y dos tipos de percepciones o actitudes del alumno, a saber:

. la motivación, el sentido de auto-eficacia y (la percepción de) el esfuerzo realizado para el aprendizaje de la Matemática, por una parte; y

. la referida a la percepción del interés y el apoyo familiar.

Además, el comportamiento de las posibles asociaciones será estudiado a la luz de otras características del alumno (género, repitencia, cambio de escuela) y de su familia (nivel socioeconómico, acceso a bienes culturales). También será determinada la posible existencia de "efecto contextual" de las variables actitudinales de los alumnos en cada escuela.

Objetivos. Más específicamente, se intentará responder a las siguientes preguntas:

1. El sentido de auto-eficacia, la motivación y (la percepción de) el esfuerzo realizado para al aprendizaje de Matemática ¿contribuyen a explicar las diferencias en el rendimiento escolar en Matemática? Además, ¿cuál es la capacidad explicativa que tiene (la percepción del alumno sobre) el apoyo e interés escolar de la familia?
2. La posible asociación de esas variables con el rendimiento, ¿se mantiene aún después de haber **controlado el nivel socioeconómico del alumno y la composición social de la escuela?** Es decir, ¿aquellas variables consiguen explicar una parte importante de la variación del rendimiento, una vez que se ha considerado la explicación proporcionada por el factor socioeconómico?
3. El "contexto escolar" relativo a cada una de las variables consideradas (auto-eficacia, motivación, esfuerzo y apoyo familiar), ¿tiene algún efecto sobre el rendimiento del alumno?
4. El efecto de cada uno de los factores hipotetizados ¿se mantiene aún cuando todos ellos actúen conjuntamente?
5. ¿Existe interacción entre el efecto de los factores estudiados y alguna característica personal del alumno?
6. La intensidad de aquellos efectos que se detecten, ¿varía de escuela en escuela?

En resumen, se busca precisar cuál de las actitudes o percepciones individuales del alumno mantiene una asociación significativa con el rendimiento escolar, mas allá del efecto que tienen el nivel socioeconómico del alumno, el contexto socioeconómico escolar y los otros factores individuales considerados.

Las variables¹². Las variables que consideraremos en este capítulo, además del rendimiento, de las que componen el "modelo de referencia" (ver Capítulo 3) y de las características personales del alumno (sexo, repitencia y cambio de escuela), son las siguientes:

(eficacia) suma del puntaje obtenido en 6 ítem tipo Likert con 4 puntos (1: muy en desacuerdo; 4: muy de acuerdo), dirigidos a medir sentido de auto-eficacia para resolver obstáculos en Matemática. Ejemplo de ítem: "Yo sé que siempre puedo resolver bien los ejercicios de matemática";

¹² Ver Cuestionario del Alumno (AP) en el Anexo.

(motivaci) suma del puntaje obtenido en 5 ítem tipo Likert con 4 puntos (1: muy en desacuerdo; 4: muy de acuerdo), dirigidos a medir motivación para el estudio de Matemática. Ejemplo de ítem: "Me gusta la matemática";

(esfuerzo) suma del puntaje obtenido en 5 ítem tipo Likert con 4 puntos (1: muy en desacuerdo; 4: muy de acuerdo), dirigidos a medir esfuerzo dedicado al estudio de la Matemática. Ejemplo de ítem: "Trato de estudiar hasta entender bien los temas de matemática".

(familia) suma del puntaje obtenido en 6 ítem tipo Likert con 4 puntos (1: casi nunca; 4: siempre), dirigidos a medir interés/apoyo de la familia en el aprendizaje del alumno. Ejemplo de ítem: "Mi familia demuestra mucho interés en mis estudios".

"Contexto" Para cada una de estas variables se construyó el indicador "contextual" correspondiente, definido como **el puntaje promedio obtenido por todos los alumnos de la escuela** en cada indicador.

Estrategia metodológica. Tal cual venimos haciéndolo, los datos serán analizados con la técnica estadística de "niveles múltiples"¹³. Para responder a las preguntas planteadas, seguiremos el siguiente camino:

7. Estimación del modelo "vacío";
8. Análisis de la asociación entre el rendimiento y cada uno de los factores específicos de este capítulo;
9. estimación del modelo de "referencia";
10. cada uno de los factores se evalúa con el modelo "referencia";
11. cada factor se evalúa con el modelo de "referencia" pero incluyendo también a los otros factores;
12. análisis de posibles interacciones entre factores;
13. finalmente, análisis de aleatorización de los factores a nivel de escuela (nivel 2);

Muestra y test de significación. El archivo de datos analizados es de 767 escuelas con 20.736 alumnos. Para evaluar la asociación entre factor y rendimiento continuaremos usando la razón de máxima verosimilitud.

Resultados. En el cuadro 5.1 se presentan los resultados obtenidos en los dos primeros pasos, con los parámetros necesarios para la comprensión del análisis (la estimación, el valor del test de máxima verosimilitud, la razón respecto al modelo "vacío" y el nivel de significación estadística).

1. El modelo "vacío". El valor del test de máxima verosimilitud arrojado por este modelo es 172957,9 (Cuadro 5.1, ítem 1). Según hemos visto en capítulos anteriores, este valor es el punto de partida para evaluar la fuerza de la relación de cada uno de los indicadores con el rendimiento.

2. Las variables individuales. Todas las asociaciones son estadísticamente significativas. En el cuadro 5.1 (parte 2) se presentan las estimaciones¹⁴, el error estándar correspondiente y el

¹³ Ver Capítulo 1 y notas metodológicas en los Capítulos 3 y 4.

¹⁴ Recordemos que las estimaciones se refieren a la pendiente de la línea que representa la relación entre cada variable y el rendimiento, es decir, la velocidad con que cambia el rendimiento cuando la variable en cuestión cambia, digamos, en una unidad.

valor del test de máxima verosimilitud. Observamos que todas las estimaciones son varias veces mayores que el error estándar correspondiente.

?? Esta conclusión se encuentra técnicamente fundamentada porque todos los tests de máxima verosimilitud arrojaron valores significativamente distantes del correspondiente al modelo "vacío", con probabilidad de ocurrencia muy inferior al 1 por mil.

Dada la definición de las variables y el signo positivo (+) de la estimación de todas esas relaciones, estos resultados permiten inferir que:

- . a mayor sentido de auto-eficacia en Matemática, más alto rendimiento;
- . a mayor motivación para el estudio de la Matemática, más alto rendimiento;
- . a mayor esfuerzo dedicado al estudio de la matemática, más alto rendimiento;
- . a mayor interés/apoyo familiar (percibido), más alto rendimiento;
- . de todas las variables consideradas, la que exhibe la mayor asociación con el rendimiento es el sentido de auto-eficacia y la de menor asociación es el interés o apoyo familiar.

Hemos respondido así a la primera pregunta planteada como objetivo al inicio de este capítulo. Nos dirigimos ahora a producir la respuesta para el segundo interrogante, a saber: ¿las relaciones detectadas, se mantienen aún después de haber **controlado el nivel socioeconómico del alumno y la composición social de la escuela?**

Cuadro 5.1 - Parámetros estimados para indicadores de actitudes y percepciones del alumno en el modelo "vacío".

Variables	Estimación	Error estándar	Verosimilitud	Nivel Signif.
1. Modelo "vacío"	-	-	172957,9	-
2. Variables individuales				
(eficacia)	0,092	(0,003)	172006,2	***
(esfuerzo)	0,065	(0,004)	172690,0	***
(motivaci)	0,043	(0,003)	172775,8	***
(familia)	0,029	(0,003)	172867,7	***

(***) : prob. ?1 por mil.

3. Modelo de referencia. El tercer paso consiste en definir el modelo que hemos denominado de "referencia. Al igual que en el caso del modelo "vacío", las estimaciones obtenidas fueron coincidentes con las exhibidas en los capítulos precedentes. Para nuestros fines, debemos retener el valor del test de máxima verosimilitud (=172226,5; Cuadro 5.2), ya que nos sirve como valor de referencia para evaluar la fuerza de la relación de cada uno de los indicadores con el rendimiento, una vez controlada la acción de las variables socioeconómicas.

4. Actitudes, percepciones y nivel socioeconómico. A continuación se evalúa individualmente cada uno de los factores con el modelo de "referencia" definido anteriormente. Las estimaciones resultantes se observan en la Parte 2 del Cuadro 5.2.

. Todas las estimaciones continúan siendo varias veces mayor que el error estándar correspondiente, es decir, la auto-eficacia, la motivación, el esfuerzo percibido y (la percepción de) el compromiso educativo familiar tienen incidencia sobre el rendimiento escolar, inclusive después de extraer la proporción de la variación explicada por las variables socioeconómicas. En otras palabras, estos factores agregan un efecto "neto", más allá del que tienen las variables del modelo de "referencia";

* Esta conclusión es confirmada por la baja probabilidad de ocurrencia de las razones de máxima verosimilitud respecto al modelo de "referencia".

. Pero por otro lado, se observa que el valor de las estimaciones disminuye respecto de lo registrado en el modelo "vacío" (Cuadro 5.1). Estas caídas evidencian que una parte del efecto inicial de todas las variables, excepto la referida a la motivación, expresaba, en realidad, la incidencia del factor socioeconómico. La medición más sensible a esa intermediación es la referida al interés o apoyo familiar percibido por el alumno. Su estimación en el modelo vacío (=0,029) desciende bruscamente en el nuevo modelo (= 0,009). Ello indica que una gran parte del efecto de la desigualdad socioeconómica sobre el rendimiento opera a través de las desigualdades en el "clima familiar educativo" que la acompaña.

Cuadro 5.2 - Parámetros estimados para indicadores de actitudes y percepciones del alumno, con el modelo de referencia.

Variables	Estimación estándar	Error militud	Verosi Signif.	Nivel
1. Modelo "referencia"	-	-	172226,5	-
2. Variables individuales				
(eficacia)	0,083	0,003	171462,5	***
(esfuerzo)	0,052	0,004	172056,3	***
(motivaci)	0,041	0,003	172062,4	***
(familia)	0,009	0,003	172218,9	0,00584

(***): prob. ?1 por mil;

5. El efecto del "contexto". En el punto anterior hemos dado respuesta al segundo interrogante planteado como objetivo de este capítulo. Ahora, abordamos la tercer pregunta, dirigida a saber la relevancia del efecto "contextual" de cada una de las variables consideradas (auto-eficacia, motivación, esfuerzo y apoyo familiar). Es decir, además del efecto de cada actitud o percepción personal del alumno, es razonable hipotetizar que el "contexto", conformado por el conjunto de las representaciones de los alumnos en el aula o en la escuela, también incide sobre el rendimiento. Se trata de probar si existe algún efecto significativo de una variable grupal, distinto y distinguible del que es propio de la variable singular correspondiente. En nuestro caso, interesa saber, por ejemplo, si además de la motivación de cada alumno, también la motivación promedio del grupo de alumnos de la escuela incide específicamente sobre el aprendizaje. De la misma forma, podemos interrogarnos acerca del sentido de auto-eficacia, del esfuerzo y del interés o apoyo familiar.

Para responder a este tipo de interrogante, a cada uno de los modelos del Cuadro 5.2. le introducimos la variable contextual correspondiente. Así por ejemplo, al modelo de referencia con la variable (motivaci), le incorporamos la variable "contextual" correspondiente, es decir

(m_motivaci) y evaluamos el comportamiento de los datos. De esta forma, podremos evaluar si las variables contextuales tienen efecto sobre el rendimiento, aún después de haber considerado el efecto de los factores socioeconómicos y de las actitudes y percepciones individuales de los alumnos. Los resultados de esta operación se encuentran en el Cuadro 5.3. A los fines de la simplicidad, presentamos solamente las estimaciones correspondiente a los términos contextuales.

Ninguno de los indicadores "contextuales" presenta indicio de efecto sobre el rendimiento. Observamos que ninguna de las estimaciones sobrepasa significativamente al error estándar correspondiente.

* El valor del test de máxima verosimilitud de cada variable contextual debe ser comparado con el de la variable individual correspondiente. Constatamos que ninguno de ellos se distancia significativamente del singular respectivo.

Estos resultados parecen sugerir que, mientras el efecto de las actitudes o percepciones del alumno individualmente considerado, no se agota ni expresa totalmente a través de la incidencia de su origen social familiar, el impacto del contexto socioeconómico de la escuela (o el aula) sí consigue expresar en su totalidad el efecto del "ethos" o cultura escolar del alumnado como grupo. Dicho de otra manera, la probable asociación del rendimiento con el "entorno actitudinal" es sólo un camino a través del cual se expresa el efecto de los factores socioeconómicos sobre el aprendizaje. Esto mismo no puede afirmarse cuando nos referimos a las actitudes o percepciones personales del alumno, al menos las que están siendo estudiadas en este capítulo.

Cuadro 5.3 - Parámetros estimados para indicadores del "contexto" de actitudes y percepciones en la escuela, con el modelo de referencia.

Variables	Estimación	Error estándar	Verosimilitud	Nivel Signif.
(m_eficacia)	0,040	(0,033)	171461,1	ns
(m_esfuerzo)	0,052	(0,048)	172055,1	ns
(m_motivaci)	0,010	(0,031)	172062,3	ns
(m_familia)	0,038	(0,040)	172218,0	ns

ns: no significativo.

El conjunto de los factores. Nos interesa ahora evaluar si el aporte individual de cada factor se mantiene aún después de considerar a los restantes. De esta forma, estaremos en condiciones de responder la cuarta pregunta objetivo de este capítulo. Para conseguirlo, deberemos incluir simultáneamente todos los factores en el modelo de referencia. Con el objeto de tornar más comprensible el comportamiento de los datos, explicitamos los resultados de acuerdo a la secuencia de modelización:

	Verosimilitud	Significación
. (eficacia)	171462,5	-
. (eficacia)+(motivaci)	171446,1	***
. (eficacia)+(motivaci)+(esfuerzo)	171445,2	ns
. (eficacia)+(motivaci)+(familia)	171437,5	0,00336

. en primer lugar, experimentamos la acción conjunta de (eficacia) y (motivaci). Aparentemente, ambas variables se mantienen como factores relevantes del rendimiento, es decir, cada uno tiene un peso importante en la explicación del rendimiento.

* La relevancia de ambas variables se comprueba porque el valor del test de máxima verosimilitud de este modelo (=171446,1) se distancia en 16,4 puntos del correspondiente al modelo que incluye sólo a (eficacia) (=171462,5), con una probabilidad de ocurrencia inferior al 1 por mil. Adviértase que aquí el nivel de significación de un modelo con 2 factores se basa en la comparación de su eficacia predictiva con la del modelo que contiene el factor más eficaz de ambos.

. A continuación, incluimos "el esfuerzo prestado al estudio de la matemática". Observemos que el valor de máxima verosimilitud obtenido se distancia levemente del modelo anterior, es decir, una vez que se han considerado la motivación y la auto-eficacia, (esfuerzo), este indicador no agrega nada importante a la explicación del rendimiento. Su efecto, anteriormente detectado, opera exclusivamente vía la motivación y el sentido de auto-eficacia. Por lo tanto, en el futuro podremos prescindir de esta variable sin perder capacidad explicativa o predictiva del aprendizaje en Matemática;

. Finalmente, sumamos al modelo la variable referida al interés y apoyo familiar percibido por el alumno. Ahora la distancia respecto al modelo anterior es indudablemente significativo. Más allá del efecto de la motivación y la auto-eficacia del alumno, (la percepción de) el apoyo e interés familiar tiene capacidad explicativa respecto del rendimiento.

Hasta aquí hemos evaluado el aporte predictivo de cada indicador cuando los otros actúan simultáneamente y de esta forma, respondimos a la cuarta pregunta establecida como objetivo de este capítulo. Sin embargo, debemos analizar ahora el comportamiento de las estimaciones de cada uno de los factores considerados. De esta forma, podremos especificar la fuerza del efecto de cada indicador. Seguiremos los mismo pasos adoptados para la presentación de los valores de máxima verosimilitud.

Al observar las estimaciones de (eficacia) y de (motivaci) cuando actúan conjuntamente, detectamos un hecho inesperado: el signo algebraico del coeficiente estimado ha cambiado de signo (de positivo a negativo), aparentando la inversión del sentido esperado de su relación con el rendimiento, de acuerdo a la estimaciones anteriores. Esto es un claro indicio de que existe una fuerte colinealidad entre ambos predictores, es decir, (eficacia) y (motivación) están altamente correlacionados entre sí y en consecuencia, no podemos especificar el efecto de cada una de ellas sobre el rendimiento.

MOTIVACI	EFICACIA
- 0,016	0,092
(0,004)	(0,004)

* Una forma de resolver el problema planteado es reducir ambas variables a una sola, a través de la creación de un nuevo indicador definido como la suma de ambas variables originales. Al experimentar esta nueva medición obtuvimos un valor de máxima verosimilitud de 172062,4, es decir, superior al conseguido con (eficacia) actuando individualmente. Por lo tanto, esta sustitución no sería eficiente.

La misma situación se produce cuando incluimos (eficacia) y (familia) simultáneamente en el modelo.

EFICACIA	FAMILIA
0,085	- 0,010
(0,003)	(0,003)

La estimación de esta última cambia de signo, indicando una correlación negativa, contrario a lo esperado, de acuerdo a las estimaciones mostradas en los Cuadros 5.1. y 5.2.

Por lo tanto, aún cuando sabemos que la motivación y el interés/apoyo familiar son factores relevantes del rendimiento del alumno, no podemos especificar su efecto cuando actúan conjuntamente con el sentido de auto-eficacia del alumno. Dado que este último indicador es el que se muestra más estable y más potente (ver Cuadro 5.2.) decidimos retenerlo para el análisis subsiguiente y prescindir de los otros dos. De hecho, él debe ser entendido en gran parte, aunque no totalmente, como un "buen representante" de la motivación y del interés familiar.

7. Características personales del alumno y la auto-eficacia. Nos proponemos ahora investigar posibles interacciones entre el sentido de auto-eficacia y las características personales del alumno estudiadas en el Capítulo 4 (género, repetencia y cambio de escuela). Esta es la quinta pregunta propuesta inicialmente para el presente capítulo.

Debemos comenzar determinando el modelo que contiene las tres características del alumno junto a las variables socioeconómicas. Las estimaciones obtenidas son:

NSE	ICUL2R2	NSE_MED	REPITE	CAMBIARO	FEMENINO
0,053	0,058	0,265	-6,114	-2,497	-1,008
(0,005)	(0,005)	(0,024)	(0,289)	(0,272)	(0,212)

Los resultados son convergentes con los alcanzados en el Capítulo 4 y por tanto, los adoptamos sin observaciones. El valor de máxima verosimilitud de este modelo es 171650,2.

A continuación, incorporamos al análisis la variable (eficacia) y obtenemos las siguientes estimaciones para las características personales del alumno:

REPITE	CAMBIARO	FEMENINO	EFICACIA
-5,574	-2,504	-0,164	0,079
(0,285)	(0,267)	(0,211)	(0,003)

El valor del test de máxima verosimilitud de este modelo es 170957,6, o sea, el efecto de (eficacia) continúa siendo significativo. Sin embargo, observamos que la estimación de (femenino) es menor que su error estándar, indicio de que ha perdido su significación. Al extraer este término y reprocesar los datos se obtiene un modelo con todos los términos estadísticamente significativos:

NSE	ICUL2R2	NSE_MED	REPITE	CAMBIARO	EFICACIA
0,044	0,040	0,297	-5,558	-2,502	0,079
(0,005)	(0,005)	(0,024)	(0,284)	(0,267)	(0,003)

El valor de verosimilitud de este modelo (=170958,3) es muy cercano al que incluía (femenino), confirmando la prescindibilidad de este último término. De esta constatación empírica podemos inferir que la posible diferencia entre géneros respecto del aprendizaje de Matemática se expresa nítidamente a través de la distribución sexuada del sentido de auto-eficacia y de la motivación para esa disciplina. Es decir, parte de esta estrecha relación entre actitudes y rendimiento acompaña la posible distribución desigual del conocimiento de la matemática entre los sexos.

Si comparamos las estimaciones de las otras variables en este modelo con las que tenían cuando no se había incluido (eficacia) veremos que todas ellas, con excepción de (nse_med), también han disminuido. Basados en este comportamiento, podemos generalizar la conclusión anterior: parte de las diferencias de rendimiento debida a desigualdades socioeconómicas, de acceso a bienes culturales y de repitencia se operan vía la distribución desigual del sentido de auto-eficacia y de la motivación entre los alumnos del sistema educativo.

Estamos ahora en condiciones de focalizar nuestra atención en las posibles interacciones de (eficacia) con otras características del alumno. Dado que el género ha perdido su relevancia, se torna interesante saber si (eficacia) tiene un nivel de incidencia distinto según sea el sexo. Para comprobarlo, creamos un término multiplicativo entre (eficacia) y (femenino) y lo incluimos en el modelo. El resultado fue negativo.

* El valor del test de máxima verosimilitud de este modelo es de 170958,3, muy próximo al anterior (=170958,3) y por lo tanto, el término interactivo no debe ser mantenido.

Por lo tanto, la fuerza de la incidencia de la auto-eficacia y la motivación sobre el rendimiento no difiere entre alumnas y alumnos.

Finalmente, interesa conocer la posible interacción entre los efectos de la composición social del grupo de alumnos o escuela, y de la auto-eficacia. La inclusión del término multiplicativo de ambas variables produjo un resultado altamente significativo. Las estimaciones de la parte fija de este modelo son las siguientes:

NSE	ICUL2R2	NSE_MED	REPITE	CAMBIARO	EFICACIA	NSE*EFIC
0,044	0,040	0,166	-5,592	-2,479	0,077	0,001
(0,005)	(0,005)	(0,032)	(0,284)	(0,267)	(0,003)	(0,000)

El signo (+) de la estimación interactiva permite deducir que el beneficio otorgado por el sentido de autoeficacia y la motivación aumenta a medida que la composición social de la escuela es más aventajada.

8. La variación del efecto sobre el rendimiento. En base a este modelo y a su valor de máxima verosimilitud (=170921,4), pasaremos a responder al último interrogante planteado como objetivo del capítulo y que ahora podemos especificar de la forma siguiente: la intensidad del efecto del sentido de auto-eficacia y de motivación para la Matemática y (la percepción de) el apoyo e interés familiar, ¿varía de escuela en escuela? Es decir, estamos interesados en saber si existen escuelas que hacen posible un rendimiento más alto que el esperado según la puntuación de los alumnos en esos indicadores actitudinales, o si, por el contrario, todas distribuyen el aprendizaje de acuerdo a lo esperado (promedio) en la totalidad del sistema.

Para responder a esta pregunta, modelaremos la variación "entre-escuela", permitiendo que el efecto de (eficacia), (motivaci) y (familia) varíen aleatoriamente en ese nivel.

* En el Capítulo 3 introdujimos el análisis de aleatoriedad de la variable intervalar (NSE) a nivel de escuela (nivel 2) (ver ecuación 8). Todas las variables que estamos analizando se asumen como intervalares y por lo tanto, se aplica aquella misma lógica.

* Adviértase que (motivaci) y (familia) ya no integran la parte fija del modelo, puesto que fueron excluidos por haber indicios de colinealidad. Sin embargo, la técnica utilizada para el análisis de los datos permite modelar la variación del efecto de cualquier variable en el nivel 1 (alumno) o 2 (escuela), aun cuando ella no tenga un efecto significativo para el conjunto de los datos (Parte Fija).

Es decir, ya no supondremos que la intensidad de asociación entre cada factor y el rendimiento es similar en todas las escuelas estudiadas. Por el contrario, abandonaremos ese supuesto, permitiendo que las relaciones varíen de escuela en escuela (aleatorización).

Como lo expresáramos al iniciar esta sección, el modelo de referencia para evaluar los resultados es al que arribamos finalmente en el punto anterior, con un valor de verosimilitud de 170921,4. Comenzamos incorporando la aleatorización del efecto del nivel socioeconómico del alumno (nse), cuya significación fuera comprobada en el Capítulo 3. A continuación, incluimos uno por vez, los indicadores bajo estudio. Los valores de los test de máxima verosimilitud de estas operaciones sucesivas para cada indicador, son los siguientes:

Indicador	Verosimilitud	Significación
Referencia	170921,4	---
+ (nse)	170891,9	***
+ (eficacia)	170854,2	***
+ (motivaci)	170848,5	0,01697
+ (familia)	170848,5	ns

Con base en estos resultados, podemos concluir que:

- ?? el efecto del sentido de auto-eficacia sobre el rendimiento varía entre las escuelas estudiadas;
- ?? también se constata una variación significativa, aunque más tenue, del efecto de la motivación sobre el rendimiento;
- ?? finalmente, el efecto del apoyo/interés familiar es constante en todas las escuelas.

El grado en que la "motivación para el estudio de la matemática" influye sobre el resultado en la prueba, varía de escuela en escuela. En algunas escuelas, el nivel de motivación o de auto-eficacia del alumno no predice el rendimiento del alumno con la misma exactitud en que muy probablemente lo hace en otras escuelas del sistema.

El Cuadro 5.4 presenta todas Las estimaciones de la parte aleatoria (nivel 2: escuela) que resultaron significativas en el análisis anteriormente expuesto.

5.4. - Estimaciones y error estandar de la Parte Aleatoria - Nivel escuela - Matemática

☺	CONS	NSE	EFICACIA	MOTIVACI
CONS	95,219 (7,106)			
NSE		0,005 (0,001)		
EFICACIA			0,001 (0,000)	
MOTIVACI				0,001 (0,000)

En resumen, los datos analizados en este capítulo permitieron constatar que el sentido de autoeficacia, la motivación y el esfuerzo realizado para aprender matemática y la percepción del apoyo familiar son variables estrechamente relacionadas con el rendimiento en Matemática. Si bien una parte importante de esa asociación expresa los condicionamientos socioeconómicos del alumno, otra (también estadísticamente significativa) refleja la incidencia específica e independiente de estas dimensiones perceptivas del alumno sobre su aprendizaje. Además, fue posible demostrar que la fuerza de la asociación entre el sentido de auto-eficacia y la motivación por un lado, y el rendimiento por el otro, varía entre las escuelas del sistema educativo.

Capítulo 6

LOS RECURSOS ESCOLARES

En este capítulo abordaremos el análisis de las relaciones entre el rendimiento del alumno y los recursos escolares. Bajo este último concepto incluiremos la existencia y el estado tanto de la infraestructura como de los recursos didácticos. La información analizada proviene de los tres cuestionarios aplicados durante el ONE de 1997 (director, docente y alumno).

Objetivos. Nos proponemos responder a las siguientes preguntas:

1. ¿Existe asociación entre los recursos escolares y el rendimiento del alumno? Más específicamente, contrastaremos la hipótesis de que a mayor disponibilidad y mejor estado de los recursos escolares, mejor será el rendimiento de los alumnos;
2. Los recursos ¿ayudan a explicar las diferencias en los rendimientos de los alumnos, aún después de haber considerado el efecto del origen social del alumno y el contexto socioeconómico y cultural de la escuela? Es decir, contrastaremos la hipótesis de que una proporción significativa de la variación dejada sin explicar por los indicadores socioeconómicos es explicada por los recursos escolares;
3. ¿Qué proporción de las diferencias de rendimiento entre las escuelas es explicada por los recursos escolares?

Estrategia metodológica. El camino seguido para el análisis está organizado en 6 etapas:

7. Estimación de los parámetros del modelo "vacío";
8. Análisis de las asociaciones iniciales entre rendimiento y cada indicador de recurso escolar;
9. Estimación de los parámetros del "modelo de referencia";
10. Análisis de las relaciones entre rendimiento y los recursos escolares, después de haber considerado el efecto del origen social del alumno y la escuela;
11. Evaluación del peso explicativo de todos los indicadores referidos a recursos, respecto a la variación en el rendimiento (Modelo final);
12. Aleatorización del efecto de la (percepción del alumno de la) infraestructura.

Los indicadores. El concepto de recurso escolar incluye la infraestructura y los recursos didácticos. Los recursos fueron mensurados de diferentes formas, siempre a través de declaraciones de los diversos actores (director, docente y alumno). En algunos casos, el respondente simplemente informa sobre la existencia o no de determinados recursos; en otros, se recoge su percepción del estado u adecuación de tales recursos. Por otra parte, algunas mediciones se refieren a toda la escuela, mientras que otras se dirigen más

específicamente al/las aula/s. Finalmente, indiquemos que las mediciones basadas en los cuestionarios de docentes se refieren al promedio que resulta de considerar a todos los cuestionarios de docentes respondidos en la escuela.

Veamos la definición de cada una de las mediciones utilizadas¹⁵:

(BIE_CULT) Debemos recordar que el índice sobre el "acceso a bienes culturales" del alumno, analizado en el Capítulo 3 anterior (Icul2r2), incluye la disponibilidad de recursos didácticos por parte del alumno. También recordemos que la asociación de este indicador con el rendimiento resultó altamente significativa y que por tanto, se lo incluyó dentro del denominado "modelo de referencia", que es utilizado a lo largo de todo el análisis para evaluar la relevancia de las diversas variables escolares. En resumen entonces, ya hemos demostrado la eficacia de la disponibilidad de recursos didácticos por el alumno para predecir su rendimiento. De aquí en más, a esta variable la llamaremos (BIE_CULT);

(DIDA_ESC) Al Director se le ofreció una lista de 20 medios y recursos didácticos y se le pidió que informase, respecto a cada uno de ellos, si existía o no en la escuela. Con base en esta información, se definió un índice sumativo, asignando 1 a cada recurso disponible. Así formada, esta medida tiene un valor máximo posible de 20 puntos;

(MAE_DIDA) En el cuestionario del docente se incluyó una medida similar a la anterior, pero con sólo 15 ítem; por lo tanto, el valor máximo del índice es de 15 puntos. La medición es el promedio de los puntajes obtenidos por todos los maestros encuestados; es decir, se trata de la disponibilidad de recursos, según el puntaje promedio de los maestros encuestados en la escuela;

(DIDA_EST) Al Director se le pidió también que evaluase el estado de los recursos disponibles, con una escala de tres valores (bueno / regular / malo). Con base en esta información se definió un índice sumativo con un valor máximo posible de 60 puntos;

(M_DID_ES) Esa misma evaluación se le pidió al docente, generándose así un índice con valor máximo posible de 45 puntos. En este caso, la medición es el promedio de los puntajes obtenidos por los maestros encuestados, es decir, se trata de la percepción promedio de los maestros acerca de las condiciones del material existente en la escuela;

(INFRA_ES) Al Director se le solicitó que evaluase el estado de 6 componentes de la infraestructura (edificio, biblioteca, aulas, baños, patios, mobiliario en general), con una escala de 3 puntos; de esta forma, quedó definido un índice sumativo con un valor máximo posible de 18 puntos;

(AULA_ESC) Además, se solicitó al Director que evaluase el estado de las mayorías de las aulas en 5 aspectos (iluminación, calefacción, ventilación, bancos, pizarrones), con una escala de 4 puntos (de muy bueno a malo); quedó definido así un índice sumativo con un máximo posible de 20 puntos;

¹⁵ Ver en Anexo el cuestionario completo del alumno (AP) y las partes pertinentes de los cuestionarios del docente (DOP) y del director (DIP).

(MAE_AULA) El docente también tuvo la oportunidad de evaluar el estado de su aula, con una escala de 4 puntos, pero sólo en tres aspectos (iluminación, calefacción, ventilación). De esta forma, definimos un índice sumativo con valor máximo posible de 12 puntos. La medición final se obtiene promediando los puntajes registrados por los maestros encuestados; es decir, se trata de la percepción promedio de los docentes acerca las condiciones físicas de las aulas;

(INFRA_AL) También se pudo conocer la percepción del alumno acerca de las condiciones de su aula, es decir, del medio ambiente físico donde se produce su aprendizaje. Fueron evaluados 6 aspectos (si era fría en invierno, calurosa en verano, luminosidad, espacio, aislamiento (ruido), mantenimiento edilicio), con dos respuestas posibles (Si/No), produciéndose un índice sumativo con un máximo posible de 6 puntos. Esta es la única medición a nivel del alumno (nivel 1) analizada en este capítulo;

(ME_INF_A) Finalmente, se estimó el promedio de los puntajes obtenidos por todos los alumnos en la medición anterior, es decir, se trata de la percepción promedio de los alumnos acerca las condiciones físicas de su aula.

La muestra. El archivo de datos analizados es de 17.991 alumnos y 677 secciones. Todas las secciones contienen 20 o más alumnos con información válida.

Los resultados. Los resultados de las tres primeras operaciones de la estrategia metodológica adoptada se encuentran sintetizados en el Cuadro 6.1.

1. El modelo "vacío". El primer paso consiste en estimar los parámetros del modelo "vacío", tal cual se ha hecho en los capítulos precedentes. Las estimaciones no difieren de las obtenidas en los modelos "vacío" anteriores. Sin embargo, debemos prestar atención al valor del test de verosimilitud (=150261,5) y a la variación "entre-escuela" (=178,389) propios de este modelo, ya que son estimaciones de referencia para evaluar la fuerza de la relación de cada uno de los indicadores con el rendimiento.

Cuadro 6.1 - Parámetros estimados para los diferentes indicadores de recursos escolares.

Indicadores	Verosimilitud	Razón de verosimilitud	Variación escuela (nivel 2)	%Disminución variación escuela
1. Modelo "vacío"	150261,5	-	178,389	-

2. Asociación inicial

Didácticos alumno

(CULT_MED)	150104,1	157,4	139,674	21,7
(BIE_CULT)	150109,5	152,0	171,316	4,0

Didácticos institución

(DIDA_ESC)	150220,8	40,7	167,411	6,2
(M_DID_ES)	150240,6	20,9	172,683	3,2
(DIDA_EST)	150245,1	16,4	173,900	2,6
(MAE_DIDA)	150248,9	12,6	174,911	1,9

Infraestructura

(MAE_AULA)	150162,9	98,6	153,069	14,2
(ME_INF_A)	150199,5	62,0	162,014	9,2
(AULA_ESC)	150206,0	55,5	163,631	8,3
(INFRA_ES)	150231,3	30,2	170,180	4,6
(INFRA_AL)	150110,1	151,4	172,914	3,1

2. Mediciones individuales. A continuación, pasamos a evaluar el grado de asociación inicial del rendimiento con cada indicador de recursos, introduciendo éstos (uno por vez) en el modelo vacío estimado en el paso anterior. Los resultados permiten constatar los siguientes hechos:

- a) todas las mediciones mantienen una estrecha relación con el rendimiento;

* La menor distancia con el valor de verosimilitud del modelo "vacío" corresponde a (MAE_DIDA) con 12,6 puntos, con una probabilidad de ocurrencia igual a 0,00038575.

b) el indicador más fuertemente asociado con el rendimiento es el referido a la disponibilidad o acceso a "bienes culturales", ya sea como variable individual (BIE_CULT) o agregada a nivel de escuela (CULT_MED). Es difícil disociar esa medición de los conceptos de "origen social" del alumno o de "composición social" de la escuela. Fue por ello que decidimos incluirlo explícitamente en el modelo que hemos denominado de "referencia" y que fuera estimado en el Capítulo 3.

Pero por otro lado, sabemos que la posesión de libros u otros materiales escolares es una variable significativamente afectada por las políticas educativas compensatorias, que incluyen su distribución a los sectores más carentes. Parece interesante entonces, contrastar su eficacia con la de los indicadores referidos a la "disponibilidad institucional" de

recursos. A ello respondió la inclusión de estas dos mediciones en el cuadro que estamos analizando;

c) De los cuatro indicadores referidos a los recursos y medios didácticos institucionales, el más eficiente es el construido en base a su disponibilidad en la escuela, según declaración del Director (DIDA_ESC). Esta conclusión parece razonable atendiendo a la mayor y mejor información del Director para responder adecuadamente. Por otro lado, la evaluación acerca del estado de los recursos no arroja resultados muy diferentes según cuál sea el origen de la misma (Director: DIDA_EST o docente: M_DID_ES). En general entonces, cuanto mayor sea la gama de recursos disponibles y mejor sea el estado de los mismos, se puede esperar que los alumnos obtengan rendimientos más altos;

d) Cuando se trata de la infraestructura, las asociaciones más fuertes se verifican con las condiciones del aula. Aún cuando las mediciones logradas con el Director son más exhaustivas (ver definición de indicadores), tanto para la escuela como para las aulas en general, son la percepción del maestro (MAE_AULA) y de los alumnos (ME_INF_A) acerca de las condiciones ambientales de sus propias aulas, las que más fuertemente se asocian con el rendimiento. Los resultados indican que cuanto mejor evaluado es el ambiente del aula, mejor es el rendimiento de los alumnos. De hecho, las condiciones ambientales de las aulas, según los maestros, explican más del 14% de la variación "entre escuelas" del rendimiento. Asimismo, en las escuelas donde la evaluación ambiental promedio de los alumnos es favorable o positiva, la probabilidad de altos rendimientos es mayor; todo lo contrario sucede en las escuelas donde sus alumnos opinan lo contrario.

e) Un párrafo aparte merece la asociación encontrada entre el rendimiento del alumno y su percepción del aula, tan fuerte como la analizada en el punto anterior.

* Recordemos que esta variable pertenece al nivel 1 de agregación, es decir, al alumno. Por ello, su efecto no es plenamente detectado por los cambios en la variación "entre-escuela" (nivel 2), sino más bien, por el valor del test de máxima verosimilitud. Se observa que su distancia con el correspondiente al modelo "vacío" (=151,4), es tan pronunciada como la detectada con la otra variable individual del alumno (BIE_CULT=152,0).

Los datos sugieren que los niños, individualmente considerados, no se conforman, adecuan o "acostumbran" a sufrir las condiciones del ambiente físico escolar, sino que, por el contrario, tienden a formarse un juicio relativo bien ajustado a la realidad.

En resumen, el comportamiento de los datos analizados confirmaron una hipótesis muy frecuente y de sentido común: **los rendimientos escolares están estrechamente asociados a la disponibilidad de recursos escolares**, al menos de la forma que fueron medidos. La nitidez con que se demuestra esa hipótesis varía. Algunas mediciones son más eficaces que otras. Cuando se trata de recursos didácticos, la posesión por parte del alumno predice el rendimiento mucho más ajustadamente que la disponibilidad en el establecimiento. Y dentro de éste último, la información del Director es más ajustada. Cuando se trata de la infraestructura, los indicadores referidos específicamente al/las aula/s son los más eficaces, sea que la información provenga de los propios alumnos o del docente.

3. El modelo de "referencia". Hasta aquí hemos analizado el grado de asociación entre indicadores referidos a recursos y el rendimiento de los alumnos, pero sin hacer intervenir las probables determinaciones sociales. Para hacerlo posible, debemos estimar el modelo de "referencia" tal cual fue definido en el Cuadro 3.3 del Capítulo 3, pero ahora en base al conjunto de datos que estamos analizando. Los resultados arrojados por esa operación (Cuadro 6.2) son muy similares a los obtenidos con anterioridad. Para los fines de nuestro análisis, sólo debemos retener el valor del test de máxima verosimilitud (=149656,1) y la variación "entre-escuela" (=121,698), ya que nos sirven como valores de referencia para evaluar la fuerza de la relación de cada uno de los indicadores con el rendimiento, una vez controlada la acción del origen social del alumno y la composición social de la escuela

4. Recursos escolares (DIP) y Nivel socioeconómico. Nos preparamos ahora para responder al 4º objetivo planteado al inicio de este capítulo, es decir, determinar si los recursos escolares mantienen una asociación significativa con la variación del rendimiento dejada sin explicar por las variables incluidas en el modelo de "referencia". No nos interesa ahora la variación total del rendimiento, sino aquella porción que no es concomitante a los indicadores de nivel socioeconómico ("residuo"). Para poder realizar esta evaluación debemos introducir cada indicador en el modelo de "referencia" y observar los cambios que ocurren. En la 2^{da} parte del Cuadro 6.2 se presentan los principales resultados.

- a) En primer lugar, notemos que el indicador referido al acceso a bienes culturales (IBIE_CULT) no aparece en esta tabla. Ello se debe a que, como lo hemos indicado anteriormente, esa medición está incluida en el propio modelo de "referencia". Esto implica que de aquí en más, las relaciones de todas las variables con el rendimiento serán "controladas" también por esa medición;
- b) En principio, ninguna de las mediciones provenientes del Director se relacionan con el rendimiento, después de haber controlado las variables socioeconómicas; éste comportamiento es indiferente al concepto que se trate de medir (disponibilidad y estado de recursos didácticos o infraestructura). Es decir, si nos hubiésemos limitado a medir los recursos a través de las declaraciones del Director del establecimiento, habríamos concluido que las relaciones inicialmente detectadas entre rendimiento y recursos son las mismas que se observan entre las variables de nivel socioeconómico y el rendimiento. El efecto de aquellas se superpone al efecto de estas últimas.

Cuadro 6.2 - Parámetros estimados para los diferentes indicadores de recursos escolares en el Modelo de "referencia".

Indicadores	Verosimilitud	Razón de verosimilitud	Variación escuela (nivel 2)	%Disminución variación escuela
-------------	---------------	------------------------	-----------------------------	--------------------------------

1. Modelo de "referencia"

149656,1 - 121,698 -

3. Asociación controlada

			Didácticos institución	
(M_DID_ES)	149651,6	4,5	120,828	0,1*
(DIDA_ESC)	149653,7	ns	--	ns
(DIDA_EST)	149656,0	ns	--	ns
(MAE_DIDA)	140652,7	ns	--	ns

Infraestructura

(ME_INF_A)	149634,1	22,0	117,560	3,4***
(MAE_AULA)	149637,5	18,6	118,210	2,9***
(AULA_ESC)	149653,4	ns	--	ns
(INFRA_ES)	149655,9	ns	--	ns
(AULA*DID)	149650,2	5,9	120,545	0,1*
(INFRA_AL)	149534,2	121,9	119,571	0,02***

(*) Prob. 5%; (***) Prob. 1 por mil.

Dado estos resultados, se decidió experimentar un término que expresa la interacción entre la evaluación del estado de las aulas y la disponibilidad de recursos didácticos, según el propio Director (AULA*DID). Los resultados obtenidos constituyen indicios de interacción entre ambos conceptos, es decir, existiría relación positiva entre rendimiento y recursos didácticos, pero solamente a medida que la infraestructura mejora. Desde este punto de vista, los efectos de ambos tipos de recursos no serían sumables sino más bien, interactivos.

- c) La superposición del efecto del nivel socioeconómico y los recursos no se mantiene cuando se consideran los otros indicadores de recursos, extraídos de las declaraciones de docentes y alumnos. Además del alto poder explicativo del indicador "acceso a bienes culturales", ya demostrado suficientemente, existe un claro indicio de que la disponibilidad y estado de los recursos didácticos en la escuela, según la evaluación del docente ((M_DID_ES), es una variable asociada al ("residuos" del) rendimiento del alumno;

- d) Respecto a la infraestructura, la observación es la misma: la evaluación que el maestro hace del estado del aula donde enseña (MAE_AULA) está fuertemente asociada con el rendimiento. Por el signo de la estimación, concluimos que cuanto mejor sea el medio ambiente del aula, se pueden esperar rendimientos más altos; o más precisamente, de dos alumnos con igual origen social y en escuelas de la misma composición social, aquél cuya aula sea más confortable tendrá mayor probabilidad de obtener alto rendimiento;

- e) Esta conclusión es aún más nítida cuando la información proviene del alumno, sea que la consideremos individual (INFRA_AL) o grupalmente (ME_INF_A). Sin duda, este comportamiento de los datos indica que aprendizaje y satisfacción con el medio ambiente inmediato donde se desarrolla la práctica pedagógica están fuertemente asociados. La valoración positiva del medio ambiente físico contribuye al aprendizaje, aún cuando se considere el peso de las diferencias sociales existentes en el sistema educativo.

5. Modelo Final Nos interesa ahora la expresión más sintética del peso total de los indicadores cuya capacidad explicativa resultara estadísticamente significativa, de acuerdo al análisis realizado hasta aquí. Para ello, permitiremos que actúen simultáneamente sobre el rendimiento, incluyéndolos en el modelo de referencia (Cuadro 6.3).

- a) Parece conveniente comenzar con el análisis de las dos mediciones extraídas del alumno, es decir, la percepción individual (INFRA_AL) y grupal (ME_INF_A) de las condiciones físicas del aula. Al hacerlo así, tendremos oportunidad de evaluar si existe algún "efecto contextual" de la percepción del medio ambiente, adicional al efecto causado por la percepción individual del alumno. Los resultados en la Parte fija del modelo son los siguientes:

CONS	NSE	BIE_CULT	NSE_MED	INFRA_AL	ME_INF_A
12,877	0,521	0,018	1,856	0,929	2,128
(2,450)	(0,034)	(0,003)	(0,163)	(0,088)	(0,652)

Observamos que el grado de asociación de ambas variables con el rendimiento continúa siendo significativo. Esto implica que no sólo el medio físico percibido por el alumno es un buen predictor de su rendimiento, sino que también éste se encuentra influido por cómo percibe el resto de sus compañeros ese mismo medio ambiente. Aparentemente, entonces, de dos alumnos con el mismo nivel socioeconómico, en escuelas de la misma composición social y con igual percepción individual del medio ambiente físico escolar, tendrá mayor probabilidad de alto rendimiento aquél que pertenezca a una grupo de alumnos que mejor califiquen a ese medio ambiente.

- b) Si a estos dos factores le agregamos la percepción del docente (MAE_AULA), observamos que la variable "contextual" (ME_INF_A) tiende a perder significación (disminuye la distancia entre estimación y error standard), mientras que la nueva variable se muestra más efectiva. En realidad, podemos pensar que estamos frente a dos indicadores que se comportan de forma muy parecida y por tanto, uno de ellos es prescindible;

CONS	NSE	BIE_CULT	NSE_MED	INFRA_AL	ME_INF_A	MAE_AULA
21,367	0,521	0,018	1,718	0,929	1,258	-0,617
(4,541)	(0,034)	(0,003)	(0,174)	(0,088)	(0,759)	(0,278)

- c) Para poder decidir, extraemos el término referido al alumno y reprocesamos el modelo. Los resultados indican que efectivamente, conviene mantener el indicador proveniente del docente y prescindir de (ME_INF_A):

CONS	NSE	BIE_CULT	NSE_MED	INFRA_AL	MAE_AULA
26,966	0,521	0,018	1,725	0,946	- 0,856
(3,041)	(0,034)	(0,003)	(0,174)	(0,088)	(0,239)

* El valor de verosimilitud de este modelo (149521,5) es inferior al que contiene a los dos términos del alumno (149523,7), indicando su mayor fuerza predictiva. Además, la distancia del valor del test de verosimilitud del modelo conteniendo los tres términos (149518,8) y el correspondiente a este modelo es igual a 2,7 que con 1 gl., tiene una probabilidad de ocurrencia por encima del 5%.

- d) Finalmente, evaluamos los dos términos restantes que resultaran significativos en el paso anterior, es decir (AULA*DID) y (M_DID_ES). Las nuevas estimaciones indican que sus efectos se superponen con los ejercidos por los indicadores anteriormente incluidos, dado que pierden significación estadística.

CONS	NSE	BIE_CULT	NSE_MED	INFRA_AL	MAE_AULA	AULA*DID
26,051	0,521	0,018	1,681	0,944	- 0,783	0,001
(3,261)	(0,034)	(0,003)	(0,184)	(0,088)	(0,257)	(0,001)

CONS	NSE	BIE_CULT	NSE_MED	INFRA_AL	MAE_AULA	M_DID_ES
25,230	0,521	0,018	1,715	0,945	- 0,793	0,071
(3,560)	(0,034)	(0,003)	(0,175)	(0,088)	(0,248)	(0,076)

El Modelo Final completo al que hemos arribado se muestra en el Cuadro 6.3.

6.3 - Modelo final de recursos escolares					
CONS	NSE	BIE_CULT	NSE_MED	INFRA_AL	MAE_AULA
26,966	0,521	0,018	1,725	0,946	-0,856
(3,041)	(0,034)	(0,003)	(0,174)	(0,088)	(0,239)
☺	CONS				
CONS	117,235				
	(6,835)				
☺	CONS				
CONS	215,065				
	(2,311)				
Test de máxima verosimilitud = 149521,5					

Esta es la forma más simple y sintética de expresar el grado de asociación entre los recursos escolares y el rendimiento. Entre el modelo de "referencia", que expresa el efecto conjunto de las mediciones sobre el origen social del alumno, y este modelo final, que adiciona el efecto de los recursos, la variación del rendimiento "entre-escuela" no-explicada ha caído en 3,7% (de 121,698 a 117,235).

* Si no nos interesase la simplicidad, podríamos considerar el efecto conjunto de las cuatro variables que han resultado significativas, sin preocuparnos del nivel de significación de cada una de ellas. En este caso, la variación "entre-escuela" cae para 116,592, es decir, el porcentaje de explicación asciende levemente a 4,1%.

Otro aspecto interesante a observar es que, al considerar la distribución y desigualdad de los recursos, el efecto de la composición social de la escuela sobre el rendimiento se ha suavizado. Esto refleja la capacidad redistributiva de la dotación de los recursos escolares; en otras palabras, estamos dimensionando la magnitud aproximada del efecto de las políticas que afectan la oferta escolar, contribuyendo a disminuir el nivel de inequidad del sistema educativo argentino.

* La disminución del peso relativo de la composición social del establecimiento se percibe en la caída de la estimación de (NSE_MED) de 2,046 en el modelo de "referencia" para 1,725 en el modelo final.

6. Aleatorización del efecto de la infraestructura. Una pregunta interesante que resta por responder es si la fuerza de la asociación entre rendimiento y (percepción de la) infraestructura es la misma (constante) en todas las escuelas o si por el contrario, ella varía significativamente. Es decir, ¿la infraestructura incide con diferente intensidad en las diversas escuelas o por el contrario, siempre tiene el mismo peso? Adviértase que este mismo interrogante fue planteado respecto de la incidencia del nivel socioeconómico del alumno (ver Paréntesis metodológico 2 y ecuación en (8) del Capítulo 3). En ese sentido,

ésta sería otra forma de evaluar la variabilidad de la "equidad educativa institucional", ahora referida a la distribución de otro bien condicionante del aprendizaje, es decir, la infraestructura y los recursos didácticos.

Adicionalmente, podríamos querer saber si existe interacción entre el nivel de rendimiento promedio obtenido por la escuela y la fuerza de aquella asociación, es decir, si a medida que aumentamos en el rendimiento promedio institucional, la incidencia de la infraestructura disminuye o aumenta.

Para responder ambas preguntas, seguimos el procedimiento de aleatorización de los parámetros vinculados a (INFRA_AL) a nivel de escuela (nivel 2)¹⁶. Los resultados se observan en el Cuadro 6.4.

6.4 - Aleatorización del efecto de (INFRA-AL) a nivel de escuela		
	CONS	INFRA_AL
CONS	133,878 (10,776) Corr: 1,000	
INFRA_AL	- 4,066 (1,532)	0,925 (0,318)

Test de máxima verosimilitud = 149508,7

El valor del test de máxima verosimilitud de este modelo es 149508,7 que, comparado con el modelo anterior (=149521,5), arroja una diferencia de 8,6. Esta diferencia, asociada a 2 g.l., tiene una probabilidad de ocurrencia de 0,013569. Aún cuando no se cumpla el criterio riguroso del 1 por mil, ese resultado es un fuerte indicio de los siguientes comportamientos:

a) la variación del efecto de (INFRA_AL) entre las escuelas, estimada en 0,925, es estadísticamente significativa;

b) a medida que descendemos en el nivel de rendimiento promedio institucional, aumenta la sensibilidad del rendimiento respecto a la infraestructura.

En resumen, **la equidad educativa (respecto a la infraestructura escolar) varía entre las escuelas**, algunas son más equitativas que otras. Y, además, a medida que las escuelas registran más bajos promedios, la sensibilidad del rendimiento a la infraestructura es mayor, es decir, son menos equitativas.

El efecto de los recursos didácticos en ONE 95. En los cuestionarios aplicados a docentes y directivos durante la evaluación de 1995 fueron incluidas también mediciones relativas a los recursos escolares del establecimiento. Esas mediciones se referían al medio ambiente físico (según el Director) y la disponibilidad de recursos didácticos (según el Director y el docente). Sin embargo, los dos indicadores que conforman el modelo final de 1997 (infraestructura percibida por el alumno y por el docente) no fueron medidos en 1995 y por

¹⁶ Ver explicación y ejemplificación de este procedimiento en el Paréntesis metodológico 2 del Capítulo 3.

lo tanto, no podemos realizar una comparación directa que permita evaluar la estabilidad de la estimación de los efectos de estos factores.

Las principales conclusiones extraídas con el análisis de los datos de 1995 fueron las siguientes:

1. En primer lugar, todas las variables mostraron una fuerte asociación inicial con el rendimiento;
2. El modelo que mejor ajustaba el rendimiento con las variables del director incluía un término interactivo entre ambiente y disponibilidad de recursos didácticos, es decir, el efecto de la disponibilidad de recursos didácticos es más evidente cuando el medio ambiente físico de la escuela es mejor. No podemos, mejorar la calidad sólo a través de la provisión de recursos didácticos, sino que debemos atender simultáneamente al contexto físico donde se desarrolla la actividad de enseñanza y aprendizaje.

CONCLUSIONES

La incidencia de los recursos institucionales (infraestructura y recursos didácticos) es confirmada ampliamente por los datos obtenidos en las evaluaciones realizadas en Argentina, mientras que la forma en que lo hacen es una cuestión abierta a investigaciones futuras.

La importancia de esta incidencia ha sido evaluada controlando la relación del rendimiento con los indicadores de nivel socioeconómico. Ello permite afirmar que los recursos escolares explican una parte significativa de las desigualdades en el aprendizaje que no es explicada por el origen social del alumno ni por la composición social de la escuela.

Se constataron diferencias en el comportamiento de los datos según cuál fuese la forma de medición utilizada y la fuente de esa información (director, docente y alumno). Parece conveniente entonces, mantener diversas mediciones de este concepto con el objeto de contar con diversas opciones y poder así evaluar adecuadamente sus relaciones con el rendimiento.

Capítulo 7

ESCUELA PUBLICA Y PRIVADA

Una de las constataciones más nítida y quizás relevante que surge de los datos de todas las evaluaciones de la calidad educativa desde 1993, es la notable diferencia entre los rendimientos escolares obtenidos por los alumnos de escuelas públicas y privadas, con evidente ventaja para estos últimos.

La reacción inmediata y más natural frente a esta constatación, es preguntar si esa diferencia en el aprendizaje se debe exclusivamente a diferencias entre ambos sectores respecto de la composición social del alumnado y/o de la dotación de recursos escolares (gasto educativo), o si, por el contrario, refleja también la incidencia de determinadas características institucionales (organización, práctica pedagógica, gestión, cultura institucional, etc.) propias del tipo de escuela privada y ausente en la pública.

Si éste último fuese el caso, entonces, es probable que algunas de las ideas bastante diseminadas en nuestra sociedad ("los alumnos de las escuelas privadas consiguen mejores aprendizajes porque la escuela privada organiza mejor, tiene una cultura organizacional más fuerte, sus docentes son más comprometidos, está mejor administrada" y así siguiendo) tenga un viso de realidad y por lo tanto, sería razonable y justificado tratar de identificar aquella/s característica/s institucional/es que explican esa mayor eficacia pedagógica para beneficio de todo el sistema educativo.

El objetivo general de este último capítulo es analizar los datos disponibles para responder con fundamento aquella pregunta general. Para una mayor precisión, podemos desdoblar ese interrogante en sus implícitos:

- (i) ¿Los alumnos de las escuelas privadas obtienen mejores rendimientos que los de la pública?
- (ii) La probable desigualdad en el nivel de rendimiento, ¿se debe exclusivamente a diferencias en el origen social del alumno y/o en la composición social del alumnado que concurre a ambos tipos de escuela?
- (iii) Además, las diferencias de rendimiento no explicadas por las variables de la pregunta anterior, ¿se deben a diferencias en la oferta o dotación (disponibilidad y estado) de recursos didácticos y/o a la infraestructura existentes en ambos tipos de escuela?
- (iv) Finalmente, interesa saber si existe alguna diferencia entre escuelas públicas y privadas, respecto de la intensidad con que el origen social (nse) y la percepción de la infraestructura por el alumno (infra_al) inciden en el rendimiento.

La primera pregunta no ofrece ninguna dificultad. Se trata de estimar los rendimientos promedios de ambos sectores y evaluar si la diferencia es estadísticamente significativa o no. La segunda y tercera son de mayor complejidad.

Desde el punto de vista lógico, estas dos preguntas deberían formularse de la siguiente forma:

?? ¿qué diferencia de rendimiento obtendría un niño con el mismo origen social que asistiese a una escuela pública y a una privada, ambas con la misma composición social?

?? ¿qué diferencia de rendimiento obtendría un niño que asistiese a una escuela pública y a una privada, ambas con la misma dotación de recursos escolares?

Obviamente, en esta situación ideal tendríamos una gran confianza al imputar la desigualdad del rendimiento a diferencias entre ambos tipos de escuela. Sin embargo, sabemos que ello no es fácticamente posible. Se requiere un "experimento" no realizable: el mismo niño debería estar simultáneamente en dos escuelas. Por eso, en su reemplazo, debemos recurrir al "control estadístico".

Cuando comparamos el rendimiento de dos alumnos, uno en la escuela pública y el otro en la privada, pero que...

... tienen el mismo puntaje en el índice de nivel socioeconómico o sea, el mismo origen social;

... son iguales también respecto a otras variables fuertemente asociadas con el rendimiento (ej. edad, repitencia, acceso a bienes culturales, etc.),

... asisten a escuelas de la misma composición social y que

... disponen de los mismos recursos didácticos e infraestructura, ...

... podemos afirmar con mayor probabilidad de certeza, que la diferencia de rendimiento entre ambos se debe a diferencias entre las escuelas, es decir, hemos "controlado" algunas variables, "igualando" a los alumnos, para poder así inferir conclusiones acerca del efecto específico del tipo de escuela.

Si pudiésemos medir todas las variables relevantes, esta estrategia nos acercaría razonablemente al modelo experimental ideal. Pero, sabemos que esto también está muy lejos de lo posible. Dado que no podemos hacerlo, siempre podrá argumentarse, por ejemplo, que no es la escuela sino más bien, una variable personal no considerada explícitamente, la que produce una determinada diferencia.

Para el tema que nos ocupa en este capítulo, este cuestionamiento metodológico tiene su expresión más evidente en la definición de la propia dicotomía público/privado. La asistencia a una escuela privada o pública es consecuencia de una decisión familiar. Esta decisión podría ser un indicador que expresa percepciones, actitudes, valores o comportamientos familiares que afectan el aprendizaje del alumno; sin embargo, existe una imposibilidad de "igualar" a los alumnos respecto a este indicador. Aún cuando el problema no tiene una solución completa, encontramos dos líneas de "contra-argumento". En primer lugar, la asistencia a una escuela privada no es materia de "decisión" para la gran mayoría de las familias, dado que no cuentan con los recursos necesarios. En segundo lugar, si bien no es posible "igualar" respecto a ese indicador específico, sí se pueden considerar otros indicadores cuyo contenido apunte en la misma dirección.

Conscientes de estas limitaciones, pasemos ahora a transformar las preguntas formuladas anteriormente en hipótesis operacionales.

1. ¿Cuál es el efecto global del tipo de escuela (pública/privada) sobre el rendimiento escolar? Es decir, ¿el rendimiento de los alumnos en escuelas públicas es diferente al de los de la privada?

2. El probable efecto que se detecte, ¿mantiene su significación estadística cuando se controla por el nivel socioeconómico de la familia del alumno (nse) y por el nivel socioeconómico promedio de la "sección" del alumno (nse_med)?
3. Si aún persiste algún efecto, ¿mantiene su significación estadística cuando se controla por el nivel socioeconómico, la disponibilidad de recursos didácticos y la infraestructura de la escuela?
4. Finalmente, ¿el efecto de (nse) e (infra_al) son igualmente homogéneos en ambos sectores (Público/Privado)?

Se indagará entonces, acerca de las relaciones entre el rendimiento escolar, el tipo de escuela (público/privado), el nivel socioeconómico y la disponibilidad de recursos escolares. Con ello, se pretende saber **si las escuelas privadas obtienen mejores resultados por sus características propiamente escolares o lo hacen por la composición social de su alumnado y la disponibilidad de recursos**. Es decir, trataremos de determinar si el tipo de escuela es propiamente un "factor escolar" o simplemente una vía a través de la cual se manifiestan las diferencias sociales subyacentes.

VARIABLES. Las variables que consideraremos son las mismas que conformaron el modelo final del capítulo anterior (nse, bie_cult, nse_med, infra_al y mae_aula), cuyas definiciones se encuentra en los Capítulos 3 y 6, y el régimen de la escuela: público y privado. Esta última variable es dicotómica y por lo tanto, deberá redefinirse como "muda" ("dummy"). Este procedimiento implica que cada alumno debe estar asignado en una y sólo en una de esas categorías¹⁷. Elegimos como "base" a la categoría "pública" y por tanto, "privada" es la variable "muda" y la denominaremos (privada).

Muestra. El archivo de datos analizados es de 17.991 alumnos y 677 secciones.

Estrategia de análisis. Para el análisis de los datos seguiremos usando, al igual que en los capítulos precedentes, la técnica de niveles múltiples. La estrategia de análisis se orienta a responder las preguntas en la misma secuencia con que fueron planteadas inicialmente. Por lo tanto, seguiremos los siguientes pasos:

1. Determinación del modelo vacío;
2. Determinación del efecto de (privada);
3. El efecto de (privada) controlando las variables socioeconómicas;
4. El efecto de (privada) controlando las variables socioeconómicas y de disponibilidad de recursos escolares;
5. Aleatorización del efecto de (nse) y (infra_al).

En resumen, esta secuencia significa que primero, se estima el efecto de (privada), y después, se lo controla por el efecto de las variables socioeconómicas y de recursos. A continuación se presentan y analizan los resultados obtenidos en cada paso.

Los resultados.

$$\begin{aligned} \text{matema}_{ij} &\sim N(X\beta, \Omega) \\ \text{matema}_{ij} &= \beta_{0ij} \text{cons} \\ \beta_{0ij} &= 51,730(0,526) + u_{ij} + e_{0ij} \end{aligned}$$

¹⁷ Ver explicación más detalladas de este procedimiento $u_{ij} \sim N(0, \Omega_u) : \Omega_u = [178,389(10,172)]$

$$e_{0ij} \sim N(0, \Omega_e) : \Omega_e = [220,996(2,375)]$$

$$-2 * \log(\text{like}) = 150261,500$$

1. Modelo "vacío". Se estiman los parámetros del modelo vacío de este archivo de datos. Estos resultados son consistentes con los obtenidos para los modelos similares en los capítulos precedentes. Aquí, sólo nos interesa retener el valor arrojado por el test de máxima verosimilitud (=150261,5).

2. Diferencia público/privado. Se incluye (privada) en el modelo vacío. Con ello, estaremos en condiciones de determinar si existe alguna diferencia significativa entre los rendimientos promedios esperados entre los alumnos de escuelas públicas y privadas. Las estimaciones para la variable (privada) indican ahora la distancia con relación a la categoría "pública". Los resultados son los siguientes:

$$\begin{aligned} \text{matema}_{ij} &\sim N(X\beta, \Omega) \\ \text{matema}_{ij} &= \beta_{0ij} \text{cons} + 10,826(1,335) \text{privada}_{ij} \\ \beta_{0ij} &= 49,889(0,551) + u_{ij} + \varepsilon_{0ij} \\ [u_{ij}] &\sim N(0, \Omega_u) : \Omega_u = [161,746(9,272)] \\ [\varepsilon_{0ij}] &\sim N(0, \Omega_\varepsilon) : \Omega_\varepsilon = [220,999(2,375)] \\ -2 * \log(\text{like}) &= 150198,800 \end{aligned}$$

La variable (privada) muestra una alta significación estadística ya que su estimación es notablemente mayor que su error standard. Se constata que el promedio de los rendimientos obtenidos por los alumnos "privados"¹⁸ es significativamente más alto (10,826 puntos porcentuales) que el de los "públicos". En este modelo, la variación "entre-escuela" del rendimiento ha caído de 178,389 en el modelo vacío, a 161,746, es decir, el 9,3% de las diferencias de rendimiento entre las escuelas se debe a la distancia entre escuelas públicas y privadas. Esta alteración es estadísticamente significativa y por lo tanto, concluimos que **los rendimientos promedios esperados en las escuelas privadas son más altos que los de las públicas.**

1. Nivel socioeconómico y escuela privada. Ahora procedemos a incluir una a una, las variables socioeconómicas y a evaluar si (privada) continúa manteniendo un efecto propio sobre el rendimiento. Los resultados son los siguientes:

Variables	Test de verosimilitud	Estimación	Error estándar	Signif.
(privada)	150198,8	10,826	(1,335)	***
+ (nse)	149819,2	8,772	(1,265)	***
+ (bie_cult)	149759,3	8,625	(1,260)	***
+ (nse_med)	149651,3	2,823	(1,284)	***

(***): prob. ?1 por mil;

Todos los valores del test de máxima verosimilitud resultantes de la inclusión secuenciada y acumulada de cada uno de los indicadores, son distantes significativamente respecto al inmediato anterior. El efecto de estos factores sobre el rendimiento no es ninguna novedad ya que había sido

¹⁸ Usamos la expresión alumno "privado" cuando nos referimos a la variación total de (matema), en contraste con la variación de (matema) en las escuelas privadas (nivel 2).

demostrado a lo largo del Capítulo 3. Más bien, nuestro interés debe centrarse ahora en el comportamiento de (privada).

Puede observarse un continuo debilitamiento de la incidencia de (privada) sobre (matema). Cuando se controla por (nse), la distancia (bruta) de rendimiento entre alumnos "públicos" y "privados" cae de 10,826 para 8,772 y a 8,625 cuando se incluye el otro indicador individual (bie_cult). Esa caída indica la incidencia del origen social del alumno que opera a través de la segmentación del sistema educativo en los sectores "público" y "privado".

Pero, es al introducir a (nse_med) cuando se produce el mayor impacto sobre la estimación de (privada), experimentando una reducción muy abrupta: la distancia entre rendimiento en ambos sectores es ahora de apenas 2,823 puntos. Tal comportamiento puede ser interpretado de la siguiente manera: **la mayor parte de la diferencia de rendimiento escolar entre alumnos "públicos" y "privados" se debe al origen social del alumno y principalmente, a la composición social de la escuela.** A pesar de ello, el efecto de la variable (privada) continúa siendo significativo.

* Para demostrar esto último realizamos el siguiente ejercicio: se extrajo (privada) de este último modelo y se calculó el test de máxima verosimilitud para las variables restantes. El resultado obtenido fue 149656,1, distante en 4.8 puntos del modelo que estamos analizando (=149651,3) que, con 1 gl., tiene una probabilidad asociada de 2,846%.

Por lo tanto, se justifica continuar con el análisis de las variables referidas a la infraestructura escolar.

4. Recursos escolares y escuela privada. Al incluir las condiciones ambientales del aula percibida por el alumno (infra_al) y por los docentes (mae_aula), la estimación del efecto de "privada" cae a 1,998. En ese momento, existe un nítido indicio de que el efecto de "privada" ha dejado de ser significativo: el valor de la estimación (=1,998) no alcanza a duplicar su error standard (1,280).

* El valor del test de máxima verosimilitud de este último modelo (=149519,0) no es significativamente distante del que se obtiene si se procesan solamente las variables socioeconómicas y de recursos escolares (=149521,5).

Las estimaciones completas de este modelo son las siguientes:

PRIVADA	NSE	BIE_CULT	NSE_MED	INFRA_AL	MAE_AULA
1,998	0,521	0,018	1,631	0,945	-0,796
(1,280)	(0,034)	(0,003)	(0,184)	(0,088)	(0,241)

☺	CONS	▲
CONS	116,763	▼
	(6,809)	
☺	CONS	▲
CONS	215,066	▼
	(2,311)	

Estos resultados indican que la diferencia público/privado en el rendimiento de los alumnos se debe exclusivamente a diferencias de tipo socioeconómicas y de recursos escolares y no a características propias de la institución escolar privada. Frente a la pregunta de cuál sería la diferencia entre ambas escuelas si la composición social y los recursos (gastos) fuesen iguales, la respuesta de los datos parece inapelable. **El rendimiento obtenido por los alumnos de las escuelas privadas no es**

superior a lo esperado de acuerdo a su origen social, a la composición social y a la dotación de recursos escolares de los establecimientos a los que asisten, según se infiere del comportamiento de la totalidad de los alumnos en la muestra.

5. Aleatorización de (nse) y (infra_al) a nivel escuela. Hasta aquí se asumió que la intensidad del efecto del nivel socioeconómico del alumno sobre su rendimiento es similar en todas las escuelas. Idéntico supuesto afectó al efecto de la (percepción por el alumno de la) infraestructura del aula. Ahora ambos supuestos serán abandonados, introduciendo su aleatoriedad a nivel de escuela (nivel 2). De esta forma, podremos registrar y comparar el grado de variación de ambos efectos en el sector público y privado.

* Para ello, se definen dos términos aleatorios en el nivel 2, uno expresando la covarianza entre (privada) y (nse), el otro, entre (privada) y (infra_al).

En primer lugar, experimentamos la aleatoriedad de (nse) respecto del tipo de escuela. Los resultados fueron negativos, es decir, no se detectó ninguna diferencia entre escuelas públicas y privada en relación a la variación de (nse) y por tanto, **ambos tipos de escuelas son homogéneas en cuanto al grado de (in)equidad educativa**, definida ésta en base al origen social del alumno.

* El test de máxima verosimilitud arrojó un valor igual a 149515,8, distante sólo a 3,7 puntos del modelo anterior, diferencia estadísticamente no significativa.

	CONS	INFRA_AL
CONS	120,595 (7,275)	
INFRA_AL		
PRIVADA		- 3,676 (1,298)
Test de máxima verosimilitud: 149511,9		

No sucedió lo mismo con el efecto de la (percepción de la) infraestructura. Los resultados obtenidos en la Parte aleatoria (nivel 2: escuela) fueron los siguientes:

* El valor del test de máxima verosimilitud de este modelo es 7,1 puntos menor que el determinado en el punto precedente, con probabilidad asociada de 0,0077 cuando se tiene 1 grado de libertad..

La significación de este término sugiere que hay diferencias entre las escuelas públicas y privadas respecto a la variación del efecto diferenciador de la infraestructura del aula. El signo de la estimación indica que entre las escuelas privadas existe una menor dispersión de las líneas de regresión que en las escuelas públicas, es decir, **el poder diferenciador de (infra_al) tiende a ser más homogéneo entre las escuelas privadas que entre las públicas.**

La estabilidad de las estimaciones a través del tiempo. Para finalizar este capítulo revisaremos brevemente los resultados obtenidos al realizar el mismo tipo de análisis con los datos de las evaluaciones de la calidad educativa de años anteriores. Dada la amplitud de las muestras y la similitud general de los indicadores utilizados, el ejercicio comparativo permite establecer conclusiones relativamente sólidas y confiables respecto del comportamiento de los datos.

El análisis de los datos de 1994¹⁹ mostró que los alumnos de 7º año de las escuelas privadas superaban en 10,64 puntos porcentuales a los alumnos de las públicas en la prueba de Matemática, cuyo promedio era alrededor de 58%. Sin embargo, cuando se controló el efecto del origen social del alumno, la composición social de la escuela y las condiciones del medio ambiente físico del establecimiento (infraestructura), esa diferencia descendió para 2,407 puntos, que con un error standard de 1,046, apenas resultó significativa al 5%.

Los datos de 1995 fueron aún más terminantes: la diferencia público/privado desaparece completamente sólo con las mediciones referidas al factor socioeconómico, sin necesidad de sumar las variables referidas a la dotación de recursos al control estadístico. En ese año, la muestra alcanzó a 18839 alumnos en 753 secciones de 7º año. Las estimaciones resultantes de incluir (privada) en el modelo vacío fueron las siguientes:

$$\begin{aligned} \text{matema}_{ij} &\sim N(\beta, \Omega) \\ \text{matema}_{ij} &= \beta_{0j} \text{cons} + 10,086(1,353) \text{privada}_{ij} \\ \beta_{0j} &= 56,160(0,567) + u_{0j} + \varepsilon_{0j} \\ [u_{0j}] &\sim N(0, \Omega_u) : \Omega_u = [192,211(10,299)] \\ [\varepsilon_{0j}] &\sim N(0, \Omega_\varepsilon) : \Omega_\varepsilon = [186,557(1,962)] \\ -2 * \log(\text{like}) &= 154433,600 \end{aligned}$$

Es decir, el rendimiento promedio de los alumnos en las escuelas privadas era 10,086 puntos porcentuales superior a los de la escuela pública. Sin embargo, cuando el efecto de (privada) se "controló" por el nivel socioeconómico del alumno (nse) y la composición social de la escuela (nse_med), esa diferencia disminuyó abruptamente, pasando a ser estadísticamente irrelevante. Las estimaciones de este último modelo son las siguientes:

$$\begin{aligned} \text{matema}_{ij} &\sim N(\beta, \Omega) \\ \text{matema}_{ij} &= \beta_{0j} \text{cons} + 1,298(1,287) \text{privada}_{ij} + 0,079(0,005) \text{nse}_{ij} + 0,339(0,026) \text{nse_med}_{ij} \\ \beta_{0j} &= 58,044(0,503) + u_{0j} + \varepsilon_{0j} \\ [u_{0j}] &\sim N(0, \Omega_u) : \Omega_u = [140,913(7,651)] \\ [\varepsilon_{0j}] &\sim N(0, \Omega_\varepsilon) : \Omega_\varepsilon = [183,722(1,932)] \\ -2 * \log(\text{like}) &= 153932,900 \end{aligned}$$

Se observa que ahora el valor de la estimación de (privada) es apenas de 1,298, con un error estándar muy próximo (= 1,287), lo cual indica claramente que la estimación no es estadísticamente significativa.

Estos resultados confirman entonces, las conclusiones extraídas anteriormente. La diferencia de rendimiento entre la escuela pública y privada se explica por diferencias en el origen social de los alumnos, en la composición social y en la dotación de recursos escolares entre ambos sectores. Los datos indican que el rendimiento de alumnos con un mismo origen social, que asisten a escuelas de igual composición social y con la misma disponibilidad de recursos escolares, no diferirá según el establecimiento sea público o privado.

¹⁹ Cervini R. y otros (1996) "Los factores del rendimiento en la educación primaria - 1994", Informe de Investigación, Ministerio de Cultura y Educación.

No parece innecesario aclarar algunos aspectos de esta conclusión. En primer lugar, se refiere estrictamente al rendimiento en Matemática. Nada se puede decir, ni menos extrapolar, respecto de otras dimensiones del resultado total del proceso de enseñanza-aprendizaje, sobre los cuales no poseemos mediciones confiables para desarrollar un análisis similar al presentado en este capítulo.

En segundo lugar, se ha excluido la dotación de recursos escolares de lo que hemos denominado "características institucionales". Con esta expresión apuntamos a las variables del "proceso educativo", pasibles de ser políticamente definidas y transformadas por **la propia institución escolar**. La dotación de recursos o "insumos educativos" está sujeta a una fuerte restricción en el sector público y no puede ser alterada significativamente a nivel de la institución escolar, a diferencia de lo que puede suceder en el sector privado. Desde el punto de vista metodológico tal exclusión está justificada dado que nuestra unidad de análisis es "la escuela", tanto pública como privada. Pero ese carácter de inmodificable no se mantiene cuando elevamos nuestra mirada hacia el sistema educativo como un todo. En este caso, sin dudas, la asignación de recursos en el sector educativo es una variable que puede ser afectada a través de decisiones políticas.

En tercer lugar, las evidencias presentadas permiten descartar la hipótesis que atribuye las diferencias de rendimiento entre ambos sectores a diferencias en sus características institucionales. Pero, de aquí no se deriva que ambos sectores sean iguales a ese respecto. Simplemente, se deduce que cualquiera sean esas diferencias institucionales, ellas no tienen capacidad explicativa en relación a la desigualdad de rendimiento entre ambos sectores. No se puede aducir entonces, que alguno de los dos sectores posea características estrictamente escolares que lo hagan más eficaz a la hora de conseguir mejores aprendizajes en Matemática.

Debería quedar claro además, que no se niega la existencia de características institucionales asociadas fuertemente a la desigualdad público/privada. Si metodológicamente renunciásemos a controlar estadísticamente esas presuntas interrelaciones por los indicadores de nivel socioeconómico y de insumo escolares, seguramente llegaríamos a una conclusión diferente a la expuesta en este trabajo. Pero al hacerlo así, estaríamos negando el acceso a un conocimiento más adecuado a la realidad. Dado que la diferencia pública / privada deja de existir, o sea, no es más estadísticamente significativa cuando se controla por los indicadores mencionados, la interpretación correcta que debe inferirse de la asociación entre la desigualdad público / privada del rendimiento y cualquier característica institucional es que ella no es más que una vía o camino a través de la cual el nivel socioeconómico y/o la dotación de recursos escolares ejercen su efecto.

Enfatizamos finalmente, que la observación anterior se refiere exclusivamente a la distancia público / privada del rendimiento y no a las desigualdades del rendimiento en general. En diversos análisis hemos demostrado que existen características de la organización y práctica institucional y de la práctica pedagógica que pueden considerarse factores del rendimiento, con efectos específicos aún después de haber controlado por los indicadores de nivel socioeconómico (individual y grupal) y la disponibilidad de recursos escolares. En efecto, las estimaciones del modelo final con los datos de 1997 (ver página 74), indican que la variación "entre-escuela" dejada sin explicar por las variables incluidas es 116,763, es decir, el 31,18% de la variación total del rendimiento. Una parte significativa de esta desigualdad podrá explicarse seguramente por características institucionales. Sin embargo el comportamiento de los datos analizados en este capítulo demuestra que tales factores institucionales se encuentran aleatoriamente distribuidos respecto de la dicotomía público / privado.