

**DETERMINACION DE FACTORES EXPLICATIVOS DE LOS
RESULTADOS ESCOLARES EN EDUCACION MEDIA EN CHILE.**

Alejandra Mizala
Pilar Romaguera

Serie Economía N° 85
Agosto, 2000

Centro de Economía Aplicada
Departamento de Ingeniería Industrial
Facultad de Ciencias Físicas y Matemáticas
Universidad de Chile

La Serie de Economía agradece el financiamiento de la Hewlett Foundation

DETERMINACION DE FACTORES EXPLICATIVOS DE LOS RESULTADOS ESCOLARES EN EDUCACION MEDIA EN CHILE.

Alejandra Mizala
Pilar Romaguera

RESUMEN

El objetivo de este estudio es indagar sobre las variables educativas, institucionales y sociales que están a la base de los logros educativos de los liceos de enseñanza media en Chile y de las diferencias de resultados entre ellos.

Utilizando datos de la prueba SIMCE de Segundo Año de Enseñanza Media del año 1998 e información del MINEDUC se estima una función de producción educacional para los establecimientos de enseñanza media del país. En primer lugar, se realiza una estimación a nivel de establecimientos educacionales utilizando el puntaje promedio en la prueba SIMCE de cada uno de ellos. En particular, se analiza la diferencia entre el puntaje esperado de cada establecimiento, dadas sus características y las de sus alumnos, y el puntaje que efectivamente obtuvieron en la prueba. En segundo lugar, se estima una función de producción educacional utilizando datos a nivel de alumnos, esta es la primera estimación de una función de producción que se realiza en Chile con datos individuales, ya que recién a partir de esta prueba SIMCE se recogió información socioeconómica de cada uno de los estudiantes que rindieron la prueba. A partir de esta estimación con datos individuales se analiza el impacto de los diferentes insumos educativos en el logro de los estudiantes.

El estudio muestra que uno de los desafíos fundamentales de la política educacional en los próximos años es reducir la brecha existente en el logro educativo entre los distintos tipos de establecimientos de enseñanza media. Es necesario lograr una mayor equidad, lo que implica que los aprendizajes deben mejorar más aceleradamente en los establecimientos que atienden a una población estudiantil de menor nivel socioeconómico. Asimismo, hay que evitar la segmentación que se tiende a producir al concentrarse los estudiantes con mayores dificultades económicas en determinados tipos de establecimientos educativos. Es importante en este punto enfatizar que el cierre de la brecha en la calidad educativa no debe implicar un descuido de aquellos establecimientos que tienen un elevado logro educativo; pues son precisamente estos liceos quienes juegan un rol fundamental en la movilidad social.

Clasificación JEL: I21, I28

Palabras claves: Economía de la Educación, Función de producción educacional

Índice

- I. Introducción
- II. Función de producción educacional
- III. Resultados de los estudios que estiman funciones de producción educacionales
- IV. Estimación de una función de producción educacional para la educación media en Chile
 - 4.1 Estimación a nivel de establecimientos educacionales
 - 4.2 Estimación a nivel de alumnos
- V. Conclusiones y comentarios finales

I. Introducción

El objetivo de este estudio es indagar sobre las variables educativas, institucionales y sociales que están a la base de los logros educativos de los liceos de enseñanza media en Chile y de las diferencias de resultados entre ellos.

Se trabaja con los resultados de la prueba SIMCE de 2° año de enseñanza media de 1998 y con una serie de variables que caracterizan a los liceos como, la dependencia, modalidad, tamaño medido por su matrícula escolar, ubicación geográfica, año de ingreso al programa MECE-Media, funcionamiento en régimen de jornada escolar completa, experiencia de los docentes, tasa alumno-profesor, si el liceo es mixto, masculino o femenino, el género del alumno, si tiene financiamiento compartido, la localidad donde está ubicado el establecimiento y características socioeconómicas de la población atendida medida a través del ingreso familiar, la educación de los padres, y el índice de vulnerabilidad de la JUNAEB.

Se trata entonces de comparar el desempeño en la prueba SIMCE de los distintos tipos de establecimientos una vez que se ha corregido por aquellas variables que no son resorte del establecimiento educacional y que sí influyen en los resultados educacionales de los alumnos. En particular se comparan los resultados educativos de los establecimientos de acuerdo a su dependencia, su modalidad y su tamaño.

Asimismo, se busca determinar el impacto de las distintas variables de caracterización de los liceos en los resultados educativos.

El trabajo está organizado en tres secciones además de esta introducción. En la primera se explica en detalle la metodología de la función de producción educacional. En la segunda sección, se resumen los resultados obtenidos en estimaciones de funciones de producción para países desarrollados, países en desarrollo y en estudios previos realizados para Chile utilizando información de las pruebas SIMCE de Enseñanza Básica. En la tercera sección se presentan los resultados de las estimaciones realizadas en este estudio para 2° Año de Enseñanza Media. En la cuarta sección se presentan las conclusiones del trabajo. En el anexo 1 se presentan correlaciones entre las distintas variables socioeconómicas utilizadas y las estadísticas descriptivas de las variables utilizadas en el estudio.

II. Función de producción educacional

La función de producción educacional es un modelo que representa simplificada el proceso educativo e intenta explicar el producto como función de variables tanto escolares como ambientales. Los estudios de la función de producción educacional son investigaciones cuantitativas de relaciones econométricas, que identifican la significancia estadística de los variados factores que influyen en el producto educativo.

La historia del análisis de la función de producción educacional comienza en EEUU, y se remonta al estudio llamado *Equality of Educational Opportunity* o más comúnmente el “*Informe Coleman*” (Coleman et. al., 1966). Aunque el Informe Coleman no fue el primer estudio de una función de producción, sí fue el más conocido y controversial, en particular porque sostuvo que los antecedentes familiares y las características de otros alumnos (*peers*) eran determinantes básicos de las diferencias en el logro educacional.

Muchos de los estudios sobre funciones de producción operan suponiendo que: las variables son conocidas con precisión por los tomadores de decisiones; la función puede ser identificada con sólo unas pocas variables de entrada (insumos); existe una relación determinística entre los insumos y el producto; y los insumos pueden ser modificados libremente por la autoridad. La realidad difiere considerablemente de las suposiciones anteriores. La función de producción es desconocida y se debe estimar usando datos imperfectos, insumos importantes no pueden ser modificados por el tomador de decisiones y muchas estimaciones de la función de producción están sujetas a considerable incertidumbre.

A pesar de las dificultades recién mencionadas, las estimaciones de funciones de producción se han convertido en un instrumento útil para examinar la importancia relativa de los diversos factores que inciden en el proceso educativo. Asimismo, las funciones de producción son un instrumento útil para hacer comparaciones de rendimiento escolar entre establecimientos educacionales que atienden a estudiantes provenientes de familias con situaciones socioeconómicas diferentes.

Se plantea el siguiente modelo, basado en la teoría microeconómica de funciones de producción.

$$L_{i,t} = f(F_{i,t}, P_{i,t}, S_{i,t}, A_{i,t}) + \varepsilon, \quad i = 1, \dots, N$$

donde:

L_i es el rendimiento escolar del estudiante i en el período t

F_i son factores acumulativos de la familia del estudiante i

P_i son características de los pares (ambiente) del estudiante i

S_i son características de la escuela y los profesores del estudiante i

A_i son características del estudiante i

ε es el error aleatorio

La medida fundamental de logro que se ha utilizado en estos estudios es test estandarizados.¹ Si bien el fin último de las escuelas queda mejor definido por variables más complejas tales como habilidad de los estudiantes para competir en el mercado laboral, mejoras de productividad, capacidad para participar en una sociedad democrática, etc.. Estas metas son virtualmente imposibles de medir contemporáneamente con la educación y sólo se pueden medir cuando ha pasado un período largo de tiempo. Es por ello que se utilizan test estandarizados como variables *proxies* de estas variables.

¹ Ver por ejemplo el programa EDURURAL en Brasil, con tests que miden desempeño en portugués y matemáticas (Harbison y Hanushek, 1992). Ver también el caso del SAT en EEUU (Power y Steelman, 1996 y Hanushek y Taylor, 1990, entre otros).

El logro de un estudiante en el instante t será el resultado de un conjunto de entradas provenientes de variadas fuentes, tales como los antecedentes de la familia del estudiante y características personales acumulados hasta el instante t , los antecedentes de los profesores del estudiante y de la escuela, acumulados hasta el instante t y otras variables como la comunidad los amigos, acumulados hasta el instante t . Estos insumos también interactúan entre sí y con las habilidades innatas o potencial de aprendizaje de los estudiantes.

Es importante puntualizar que el proceso educacional debe ser visto como acumulativo; algunos insumos pasados tienen efectos duraderos, en tanto otros pueden disminuir con el tiempo².

Los insumos familiares tienden a ser medidos por características sociodemográficas, tal como educación de los padres, ingresos y tamaño familiar. Insumos de los pares cuando se incluyen, son típicamente resumidos a través de características sociodemográficas de otros estudiantes en la escuela. Los insumos de la escuela incluyen características de los profesores como nivel de educación, experiencia, sexo, etc.; en cuanto a la organización de la escuela se toman variables tales como: tamaño de la clase, facilidades, gastos administrativos, etc.

Los problemas más frecuentes de estas estimaciones son dos: primero, no existe una adecuada medida de habilidad de los estudiantes; segundo, mientras la educación es acumulativa, frecuentemente sólo existen medidas contemporáneas de los insumos. Cada uno de estos problemas conduce a un sesgo en los efectos estimados de los insumos educacionales.

Es natural que muchos factores de los estudiantes no puedan ser adecuadamente medidos, el punto clave es si el conjunto de estos factores se relaciona con las observaciones de la familia, escuelas y otras influencias sobre el logro educacional. Si las variables omitidas no están correlacionadas con las variables explicativas, las regresiones arrojarán estimaciones ineficientes pero insesgadas; en cambio, si están correlacionadas las estimaciones serán sesgadas e inconsistentes y no podrán utilizarse para la evaluación o análisis de política.

El término de error ε probablemente contiene una variedad de factores no medidos. Primero, la educación es un proceso acumulativo, el pasado histórico completo es necesario para caracterizar el logro educacional en cualquier instante del tiempo. Segundo, la mayoría de los estudios limitan el rango y características de las observaciones; por ejemplo, aún con las pruebas diseñadas especialmente, es difícil registrar las diferencias cualitativas de las conductas de los profesores. Tercero, muchos de los factores son difícilmente medibles; por ejemplo, la mayoría de los estudios muestran que las diferencias en las habilidades innatas de los estudiantes son importantes en la determinación del logro escolar, pero pocas veces especifican la forma de medir las habilidades innatas, similarmente, la motivación y aspiraciones de los estudiantes son extraordinariamente difíciles de cuantificar.

² Para más detalle ver Hanushek, 1986

Una forma de abordar estos problemas es reformular el modelo del logro, mirando el incremento en el tiempo. Si por ejemplo se puede observar el logro al final de un instante t^* , es posible analizar $(L_t - L_{t^*})$. Intuitivamente el incremento en el desempeño dice relación con los factores de entrada como profesores y escuelas entre el período entre t^* y t , en tanto que la estimación controlaría por factores comunes en ambos períodos, como aquellos que están asociados a las características de los alumnos y sus familias. Esta formulación sería una mejor representación de un *modelo de valor agregado*, la cual requiere contar con observaciones en dos momentos del tiempo para el mismo grupo de estudiantes.

En los estudios empíricos L_{t^*} es frecuentemente incluida como una variable explicativa en la regresión (en vez de considerar como variable dependiente el cambio en el logro). Esto tiene dos ventajas. Primero, permite diferenciar el crecimiento en el logro por sobre un puntaje inicial. Segundo, permite utilizar diferentes unidades de medida de logro a través del tiempo, por ejemplo, usar dos test distintos en años diferentes.

Otros problemas que se enfrentan al estimar funciones de producción son los siguientes. En primer lugar, la especificación apropiada de la función de producción educacional, incluyendo su forma funcional, el nivel de agregación y las variables de control relevantes es incierta Krueger (1997). En segundo lugar, resultados poco satisfactorios pueden ser explicados por la presencia de ineficiencias en los datos, es decir, ineficiencia en el proceso productivo educacional. Esto significa que dado los recursos con que cuentan los establecimientos éstos no están obteniendo su producto potencial; por lo tanto, no estamos estimando la frontera de producción sino un punto al interior de ésta. Esta ineficiencia tiene un carácter técnico y no se la debe interpretar como ineficiencia en la asignación de los recursos (Deller y Rudnicki, 1993; Mizala, Romaguera y Farren, 1998).

III. Resultados de los estudios que estiman funciones de producción educacionales

Uno de los elementos que todos los estudios consideran son las variables familiares. Berger y Toma (1994), Deller y Rudnicki (1993), Hanushek y Taylor (1990), Summers y Wolfe (1977), entre otros, encuentran que la educación de los padres es una variable estadísticamente significativa para explicar el desempeño de los niños; no obstante, los resultados respecto del tamaño de la familia son ambiguos. Asimismo, el ingreso familiar es una variable muy importante. Este es un resultado que se confirma en todos los estudios de funciones de producción educacionales.

Respecto de las características de los estudiantes, la investigación de Summers y Wolfe (1977), muestra que el sexo del estudiante afecta el aprendizaje, los hombres tienen en promedio peor rendimiento que las mujeres. La motivación de los estudiantes (medida por las ausencias no justificadas y por atrasos), tiene también efectos significativos sobre el aprendizaje, los estudiantes con más ausencias sin justificar tienen un menor crecimiento en el logro.

Summers y Wolfe (1977), al estudiar el efecto de las variables relacionadas con la escuela, encuentran que el tamaño de los cursos no es significativo para explicar el crecimiento en el logro y que, al parecer, los alumnos matriculados en escuelas más pequeñas tienen mejores resultados que los alumnos de escuelas grandes, principalmente en el caso de los estudiantes de raza negra. Sin embargo, Krueger (1997) utilizando un experimento realizado en el proyecto STAR en EE.UU., donde los estudiantes son asignados aleatoriamente a clases con diferentes insumos educativos, muestra que alumnos que asisten a cursos pequeños obtienen mejores resultados en pruebas estandarizadas. En cuanto al número de libros por alumno, Summers y Wolfe (1977) muestran que existen efectos perversos; es decir, el número de libros por alumnos tiene un efecto negativo. Por su parte, las características físicas de las escuelas parecen no hacer mucha diferencia en cuanto al logro. Deller y Rudnicki (1993), también con relación al impacto de la escuela, encuentran que gastos en docencia tienen efectos significativos en el mejoramiento del desempeño escolar y, al contrario, un aumento en los gastos administrativos es un obstáculo para un buen desempeño de los estudiantes. Adicionalmente, escuelas de mayor tamaño tienen menor rendimiento que escuelas con menos alumnos.

Con relación al impacto que tiene sobre el logro escolar las características de los profesores, Berger y Toma (1994), utilizando un modelo de valor agregado, encuentran que aquellas escuelas que tienen como requerimiento el grado de master para la certificación de los profesores, tienen un rendimiento más bajo. Por su parte, el salario de los profesores no es significativo para explicar el logro de los alumnos. Summers y Wolfe (1977), también encuentran que la educación de los profesores no parece estar positivamente correlacionada con un mayor aprendizaje de los estudiantes. Este mismo estudio concluye que la experiencia de los profesores no es significativa y que existe una relación perversa entre los resultados de exámenes que rinden los profesores (National Teacher Exam Score³) y el aprendizaje de los alumnos; no obstante, afirman que los resultados de este examen no pueden ser usados como única medida del potencial del profesor. En cuanto a las variables que aproximan la calidad de los profesores, como son la experiencia y el salario, ninguna resulta estadísticamente significativa en la explicación del logro de los alumnos.

El estudio de Berger y Toma (1994), investiga también el efecto sobre los resultados educacionales del gasto por alumno que realiza cada Estado en EE.UU., las estimaciones muestran que aquellos Estados que gastan más por alumno tienen rendimientos significativamente más bajos.

La mayoría de las investigaciones realizadas para países en desarrollo (PED) han sido revisadas en un trabajo de Fuller y Clarke (1994). Ellos concluyen que en los PED no se puede asumir que las características familiares es lo único que importa y que, por tanto, el colegio tendría poco que ver con los resultados de los alumnos. Al contrario, existe mucha evidencia empírica mostrando que hay insumos del proceso de producción educacional que pueden ser modificados para mejorar los resultados; por ejemplo, el tiempo de instrucción, la disponibilidad de textos y material de lectura y ciertos métodos de enseñanza. En particular, Harbison y Hanushek (1992) en su estudio sobre Brasil reportan que la infraestructura y los materiales de aprendizaje de la escuela, están positivamente

³ El National Teacher Exam es una prueba que se les aplica a los profesores en EEUU.

correlacionados con el desempeño de los alumnos, variables tales como servicios básicos y características de las salas de clases influyen en forma significativa en el logro escolar.

Asimismo, Mizala y Romaguera (1999) examinan los factores que afectan el desempeño escolar en Bolivia, para ello utilizan los resultados obtenidos en las primeras pruebas de medición de la calidad de la educación aplicada en ese país. Las estimaciones muestran que las variables socioeconómicas y raciales son muy importantes para explicar el desempeño escolar; no obstante, también se encuentra que las variables asociadas con las escuelas y los profesores (experiencia de los profesores, tareas diarias para la casa, tamaño de la escuela) son muy relevantes. También concluyen que las escuelas públicas tienen peores resultados que las privadas, una vez que se ha controlado por otros factores que afectan el desempeño escolar.

Por su parte, Hanushek (1995) revisa 96 estudios que estiman los efectos de los insumos educacionales en el desempeño de los estudiantes en PED. Los insumos considerados son la tasa alumno profesor, la educación, la experiencia y el salario de los profesores, el gasto por alumno y la infraestructura y recursos del establecimiento. Los resultados obtenidos no apoyan la idea de que cursos más pequeños, profesores con mayor experiencia y mejores salarios tienen un efecto positivo sobre el desempeño estudiantil. Adicionalmente, la educación de los profesores y el gasto por alumno tienen un efecto positivo y significativo en alrededor de la mitad de los estudios revisados. La evidencia empírica muestra que los insumos más importantes, en términos del desempeño de los alumnos, son la infraestructura y los recursos con que cuenta el establecimiento, es decir, disponibilidad de textos, ciertos métodos de enseñanza y tiempo de instrucción entre otros (65% de los estudios reportan un coeficiente positivo y estadísticamente significativo en una ecuación de logro).

Las investigaciones realizadas en Chile previamente difieren de los demás estudios, debido principalmente a que la unidad de análisis son las escuelas y no los alumnos como en los trabajos anteriores, ya que no se contaba con datos a nivel individual de las características de las familias.

Los estudios para Chile utilizan los resultados obtenidos por los establecimientos educacionales en las pruebas SIMCE de enseñanza básica. Estos son: Rodríguez (1988), Aedo y Larrañaga (1994), Aedo (1997), Carnoy y McEwan (1997), Bravo, Contreras y Sanhueza (1999) y Mizala y Romaguera (1998) y (2000). En general estos trabajos concluyen que las características socioeconómicas de las familias son variables estadísticamente significativas para explicar el desempeño de los estudiantes en los diferentes tipos de establecimientos educacionales.

Sin embargo, sus conclusiones difieren al comparar el desempeño de los establecimientos de acuerdo a su dependencia. Rodríguez (1988), Aedo y Larrañaga (1994), y Aedo (1997) concluyen que los establecimientos particulares subvencionados tienen mejores resultados en las pruebas SIMCE que los establecimientos municipales, mientras que Carnoy y McEwan (1997) encuentran que los establecimientos municipales tienen un mejor desempeño que los particulares subvencionados. Los estudios difieren básicamente en las variables de control utilizadas y en las muestras de establecimientos con que trabajan. Rodríguez (1988), Aedo y Larrañaga (1994), Aedo (1997) se basan en pequeñas muestras

(menos de 1000 establecimientos), lo que no les permite generalizar sus resultados. Carnoy y McEwan (1997), inicialmente utilizan todo el universo de establecimientos que rinden la prueba SIMCE, no obstante, luego reducen la muestra y excluyen a los establecimientos particulares pagados al incorporar nuevas variables de control. Mizala y Romaguera (1998) y (2000) trabajan con todos los establecimientos que rindieron la prueba SIMCE de 4° Básico en 1996 y concluyen que no hay diferencias estadísticamente significativas en el desempeño de los establecimientos municipales y los particulares subvencionados en ese año, las diferencias se producen entre éstos y los particulares pagados. Este estudio concluye que la muestra de establecimientos con que se trabaja puede sesgar los resultados y por ello es necesario utilizar el universo de establecimientos que rinden la prueba. Bravo, Contreras y Sanhueza (1999) generalizan los resultados de Mizala y Romaguera (1998) y (2000) estimando funciones de producción educacionales para los diferentes años y cursos (4° y 8° Básico) que han rendido la prueba SIMCE, concluyendo que no se observan diferencias entre los establecimientos particulares subvencionados y los municipales, una vez que se controla por las otras variables que influyen en el logro educativo, principalmente las características socioeconómicas de las familias.

Al revisar los estudios empíricos sobre la función de producción educacional en PD llama la atención que en varios de ellos se concluye que algunos de los insumos tradicionales no tienen un efecto estadísticamente significativo sobre el logro de los estudiantes. Sin embargo, esto no ocurre necesariamente en los estudios para PED, posiblemente porque en estos países se observa una mayor dispersión en las características de los establecimientos educacionales y los estudiantes, lo que permite estimar mejor la influencia de estos factores sobre el logro educativo.

IV. Estimación de una función de producción educacional para la educación media en Chile

Utilizando datos de la prueba SIMCE de Segundo Año de Enseñanza Media del año 1998 e información del MINEDUC se estima una función de producción educacional para los establecimientos de enseñanza media del país. En primer lugar, se realiza una estimación a nivel de establecimientos educacionales utilizando el puntaje promedio en la prueba SIMCE de cada uno de ellos; en particular, se analiza la diferencia entre el puntaje esperado de cada establecimiento, dadas sus características y las de sus alumnos, y el puntaje que efectivamente obtuvieron en la prueba. En segundo lugar, se estima una función de producción educacional utilizando datos a nivel de alumnos, esta es la primera estimación de una función de producción que se realiza en Chile con datos individuales, ya que recién a partir de esta prueba SIMCE se recogió información socioeconómica de cada uno de los estudiantes que rindieron la prueba. A partir de esta estimación con datos individuales se analiza el impacto de los diferentes insumos educativos en el logro de los estudiantes.

La prueba SIMCE que se aplicó en 1998 difiere metodológicamente de las anteriores pruebas en que tiene una escala no acotada que mide habilidades de los estudiantes, la anterior prueba tenía un puntaje que fluctuaba entre 0 y 100. En esta nueva prueba se relaciona el puntaje de los alumnos con sus habilidades, incluyendo distintos tipos de

preguntas asociadas a diferentes habilidades⁴. Para realizar la comparación entre distintos niños lo que se hace es incluir items (preguntas) que superan las habilidades del más hábil y a partir de éste se ordenan (rankean) los estudiantes. El puntaje promedio se asigna de acuerdo a la dificultad media de la prueba. Al mismo tiempo esta prueba contiene items comunes con pruebas de otros años, lo que permite corregir los resultados de forma de hacerlos comparables en el tiempo. La prueba tiene una media de 250 puntos con una desviación estándar de 50 puntos.

En este sentido, a diferencia de lo que ocurría con las anteriores pruebas SIMCE, esta nueva metodología permite a los establecimientos mejorar constantemente. En el caso anterior era muy difícil mejorar para un establecimiento que tenía un puntaje superior a los 90 puntos.

4.1 Estimación a nivel de establecimientos educacionales

En primer lugar, se intentó una estimación a nivel de establecimientos, utilizando datos promedio de los establecimientos educacionales. Sin embargo, esta regresión adolecía de multicolinealidad, fundamentalmente entre la variable dependencia del establecimiento y los indicadores socioeconómicos. El problema de multicolinealidad persistió aún utilizando alternativamente la educación de los padres, el ingreso familiar y un índice socioeconómico construido a partir de estas dos variables.

El índice socioeconómico se construyó utilizando análisis factorial. Lo que se hace es obtener un promedio ponderado de ambas variables (educación e ingreso familiar), donde los ponderadores se obtienen de maximizar una función de verosimilitud, de esta forma éstos permiten capturar la mayor información (varianza) de ambas variables. En el anexo se presentan las correlaciones entre el índice y las variables que lo componen (cuadro A1).

El cuadro 1 muestra la fuerte correlación existente entre las variables socioeconómicas y la dependencia del establecimiento educacional⁵.

⁴ En este sentido un niño puede tener 5 preguntas buenas y obtener 250 puntos en la prueba y otro niño tener otras 5 preguntas buenas y obtener 300 puntos, ya que las preguntas miden diferentes grados de habilidad.

⁵ Hay que hacer notar que estos problemas no se presentan o son mucho menos significativos en 4° Básico, lo que puede ser una indicación de que hay una mayor segmentación socioeconómica en enseñanza media.

Cuadro 1. Correlación entre dependencia y variables socioeconómicas

	Dependencia
Índice socioeconómico	0.800** (1661)
Educación de los padres	0.782** (1662)
Ingreso familiar	0.759** (1661)

Notas: (1) ** la correlación es significativa al 1% (bilateral).

(2) El número de establecimientos está entre paréntesis

(3) Las dependencias son: municipal, particular subvencionado y particular pagado

Por esta razón a nivel de establecimientos educacionales se optó por estimar una regresión en que la variable a explicar es el resultado del establecimiento en la prueba SIMCE de 2° año de E. Media en 1998, pero de la cual se omitió la dependencia del establecimiento como variable explicativa⁶.

A partir de los resultados de esta regresión se obtuvo el puntaje esperado de cada establecimiento dadas las características de éste y de sus alumnos. Es decir, el puntaje esperado lo da el modelo estimado. Luego, se calculó la diferencia entre el puntaje efectivo del establecimiento y su puntaje esperado. Esta medida nos indica la brecha existente entre el puntaje que efectivamente obtiene el establecimiento y el que debiera obtener dadas sus características y las características de la población que atiende. Esta diferencia entre puntaje efectivo y puntaje esperado se analizó para cada una de las dependencias de los establecimientos. Si un establecimiento tiene una diferencia negativa significa que su puntaje efectivo es menor que el esperado dado por el modelo.

La estimación se realizó por separado para las pruebas de castellano y de matemáticas. La regresión estimada tanto para castellano como para matemáticas se presenta en los cuadros A2 y A3 del anexo.

El cuadro 2 presenta un resumen de las estadísticas descriptivas de la brecha entre puntajes efectivos y esperados. Los gráficos 1 y 2 muestran los rangos y la distribución de la brecha de puntaje, diferenciando por dependencia y para las pruebas de castellano y matemáticas respectivamente.

⁶ Por razones de multicolinealidad también se omitió el uso del Índice de Vulnerabilidad de la JUNAEB.

Cuadro 2. Estadísticas sobre las diferencias entre puntaje efectivo y puntaje esperado

	Media	Desviación típica	N (número de establecimientos)
Castellano			
Municipales	-2.0	16.7	545
Particular Subvencionados	2.2	19.1	618
Particular Pagado	-0.9	22.8	332
Total	0.0	19.3	1495
Matemáticas			
Municipales	-1.9	16.2	545
Particular Subvencionados	2.2	20.2	618
Particular Pagado	-1.0	25.5	332
Total	0.0	20.3	1495

Nota: El puntaje esperado se estimó a partir de una función de producción con las siguientes variables dependientes: Índice de nivel socioeconómico: construido a partir de la educación de los padres y el ingreso familiar; índice de nivel socioeconómico al cuadrado; modalidad: Científico Humanista, Técnico Profesional y Polivalente; experiencia promedio de los docentes del establecimiento; tasa alumno/profesor de aula; matrícula de enseñanza media del establecimiento; matrícula al cuadrado; género: mujeres, hombres, mixto año de ingreso al MECE: 1994, 1995, 1996, 1997, 1998 o más; financiamiento compartido: 0 si no tiene o el monto correspondiente si tiene; si el establecimiento está o no en régimen de jornada completa⁷.

Llama la atención que los establecimientos particulares pagados tengan una media menor y una mayor desviación típica que los establecimientos particulares subvencionados. Esto puede significar que hay establecimientos particulares pagados que tienen, tanto en castellano como en matemáticas,⁸ un puntaje efectivo menor que el que se esperaría de sus alumnos dadas las características de éstos y del establecimiento educacional.

Adicionalmente, las medias de los establecimiento municipales son más bajas, a pesar de que el límite izquierdo de la distribución no es menor; la mayor brecha negativa para los municipales (matemáticas: -50) es menor que la de los particulares subvencionados (matemáticas: -70). El hecho que la media sea más pequeña para los establecimientos

⁷ Otras variables utilizadas que no resultaron estadísticamente significativas fueron: si el establecimiento es urbano o rural y diversas variables que intentaban medir el grado en que los habitantes de una comuna tenían diversas opciones donde matricular a sus niños (particular pagado, particular subvencionado y municipales). El que estas variables no resulten significativas tiene relación con las dificultades existentes para medirlas adecuadamente y no necesariamente se puede concluir que éstas no afectan los resultados en el SIMCE de los establecimientos.

⁸ Hay un 27% de establecimientos particulares pagados para los cuales no dispusimos de información, el porcentaje de establecimientos municipales sin información es 6% y un 16% de los particulares subvencionados.

municipales está influida por una mayor concentración de establecimientos con puntajes ligeramente negativos (matemáticas: -10) y por el bajo número de establecimientos de este tipo que tienen una brecha alta y positiva entre el puntaje efectivo y el puntaje esperado.

No obstante es interesante, más allá de observar los gráficos, determinar si estadísticamente las distribuciones de las diferencias entre los resultados esperados y los efectivos son iguales para cada uno de los tipos de establecimientos educacionales. Para ello utilizamos test no paramétricos ya que éstos tienen la ventaja sobre los test paramétricos que no requieren suponer a priori una función de distribución determinada.

Utilizamos tres tests que permiten contrastar estadísticamente las formas de las distribuciones: el test U de Mann-Whitney, el test de Wilcoxon y el test de Kolmogorov-Smirnov. El primero de éstos es la más conocida de las pruebas para dos muestras independientes. El test de Mann-Whitney contrasta si dos poblaciones muestradas son equivalentes en su posición, el test se basa en los signos (positivos o negativos) de los residuos de la regresión. El segundo test no sólo se focaliza en el signo, sino en el monto de la diferencia entre el puntaje efectivo y el puntaje esperado de cada establecimiento educacional entregado por la regresión. Los estadístico U de Mann-Whitney y W de Wilcoxon están relacionados ya que suman una constante asociada al tamaño de los grupos, de manera que el estadígrafo Z (que es la normalización de U) es idéntico para U y W. Este estadígrafo Z tiene una distribución asintótica normal estándar (N(0,1)). La hipótesis alternativa de esta prueba establece que las distribuciones no son las mismas, pero implica solamente un desplazamiento en la tendencia central de una con respecto de la otra y no una diferencia de forma o dispersión.

El tercer test de Kolmogorov-Smirnov es más general que los anteriores y detecta las diferencias entre las posiciones y formas de las distribuciones. Se basa en la diferencia máxima absoluta entre las funciones de distribución acumulada observadas para ambas muestras, en definitiva prueba si las dos muestras analizadas provienen de la misma distribución poblacional.

Los tres test permiten comparar las distribuciones de a pares de muestras, los cuadros que presentamos a continuación resumen los resultados de la aplicación de estos tests.

Cuadro 3 : Establecimientos municipales versus particulares subvencionados

	Prueba de castellano	Prueba de matemáticas
U de Mann-Whitney	140765.0	144878.0
W de Wilcoxon	289550.0	293663.0
Z (significación asintótica)	-4.836** (0.000)	-4.116** (0.000)
Kolmogorov-Smirnov (significación asintótica)	2.533** 0.000	2.313** 0.000

Nota: ** Estadísticamente significativo al 1%

Los resultados del cuadro 3 muestran que las distribuciones de las diferencias entre los puntajes esperados y los efectivos son diferentes entre los establecimientos municipales y los particulares subvencionados. Lo mismo ocurre con las diferencias entre los establecimientos municipales y los particulares pagados (ver cuadro 4).

Cuadro 4 : Establecimientos municipales versus particulares pagados

	Prueba de castellano	Prueba de matemáticas
U de Mann-Whitney	79066.0	79754.5
W de Wilcoxon	227851.0	228539.5
Z (sig. asintótica)	-3.134** (0.002)	-2.945** (0.003)
Kolmogorov-Smirnov (sig. asintótica)	2.185** (0.000)	2.353** (0.000)

Nota: ** Estadísticamente significativo al 1%

No obstante, al comparar las distribuciones de los establecimientos particulares pagados y subvencionados no se puede rechazar la hipótesis de que ambas son diferentes (ver cuadro 5). Esto es, los resultados obtenidos indican que la brecha entre los puntajes esperados y efectivos de los establecimientos particulares subvencionados tiene una distribución similar a la de los establecimientos particulares pagados.

Cuadro 5: Establecimientos particulares subvencionados versus particulares pagados

	Prueba de castellano	Prueba de matemáticas
U de Mann-Whitney	97265.0	100890.0
W de Wilcoxon	152543.0	156168.0
Z (sig. asintótica)	-1.320 (0.187)	-0.421 (0.674)
Kolmogorov-Smirnov (sig. asintótica)	0.990 (0.281)	1.078 (0.196)

Nota: ** Estadísticamente significativo al 1%

Este resultado es consistente con el análisis realizado por Bellei (1999) acerca de la evolución de los puntajes obtenidos por los establecimientos educacionales entre el SIMCE de 2° Medio de 1994 y el de 1998. De acuerdo a éste se observa que los establecimientos particulares subvencionados experimentaron una significativa mejora, sobretudo en la prueba de matemáticas.

GRAFICO 1

GRAFICO 2

4.2 Estimación a nivel de alumnos

A continuación se estima una función de producción educacional a nivel de alumnos, la variable a explicar es el logro individual del alumno en las pruebas de castellano y de matemáticas en la prueba SIMCE de 1998. Las variables explicativas corresponden a variables asociadas a los alumnos y a variables asociadas a los establecimientos educacionales⁹, estas son¹⁰:

- Índice socioeconómico del alumno, construido con sus datos de ingreso familiar y educación de los padres¹¹
- Índice socioeconómico del alumno al cuadrado
- Género del alumno: hombre o mujer
- Dependencia: municipal, particular subvencionado, particular pagado
- Modalidad: Científico Humanista, Técnico Profesional y Polivalente.
- Experiencia promedio de los docentes del establecimiento
- Tasa alumno/profesor de aula
- Matrícula de enseñanza media del establecimiento
- Matrícula al cuadrado
- Año de ingreso al MECE: 1994, 1995, 1996, 1997, 1998 o más
- Financiamiento compartido: 0 si no tiene y el monto correspondiente si tiene
- Si el establecimiento está o no en régimen de jornada completa

Tal como se hizo en la estimación a nivel de establecimientos educacionales, se incluyen la variable nivel socioeconómico al cuadrado y matrícula al cuadrado para investigar si hay algún grado de no linealidad (concavidad) en la relación entre estas variables y los resultados del SIMCE.

Es importante comentar que hay variables que pueden incidir en el resultado escolar de los alumnos y que; sin embargo, no tenemos posibilidades de medirlas, este es el caso las

⁹ En Chile no es posible estimar una función de producción educacional en valor agregado, ya que debido a la calendarización del SIMCE no se le toman pruebas a la misma generación de alumnos en dos momentos del tiempo. Una posibilidad es incluir como variable explicativa en la regresión a nivel de establecimientos, los resultados obtenidos por el mismo liceo en la prueba de Segundo Año de Enseñanza Media del año 1994; sin embargo, esto no se puede realizar en las regresiones a nivel de alumnos. Esta es la primera estimación de una función de producción educacional realizada con datos individuales en Chile.

¹⁰ En el anexo se presentan las estadísticas descriptivas de estas variables tanto para la muestra de alumnos de la prueba de castellano como la de matemáticas (cuadros A5 y A6). Sólo 21 niños pertenecían a establecimientos con menos de 20 alumnos, por lo que no fue necesario ponderar los datos. Las variables “proxy” que intentan medir la diversidad de la oferta educativa en la comuna no resultaron significativas por lo que no se incluyeron en la regresión.

¹¹ En el cuadro A4 del anexo se presentan las correlaciones entre el índice socioeconómico y las variables ingreso de la familia y educación de los padres. La metodología seguida es la misma que se explicó más arriba. Para la variable ingreso familiar se utilizó el punto medio del intervalo escogido por los padres en la encuesta.

características intelectuales de los estudiantes, la capacidad de los profesores expresadas en prácticas pedagógicas (sólo estamos midiendo años de servicio), y los recursos con que cuenta cada establecimiento educacional¹². Estas variables pueden ser más importantes al realizar estimaciones de logro educativo a nivel de estudiantes que estimaciones a nivel de establecimientos escolares. Esto podría explicarse por que la bondad de ajuste de los modelos (R^2 ajustado) es menor para datos individuales que para datos a nivel de liceos¹³.

Los resultados se presentan en el cuadro 6 para la prueba de castellano y en el cuadro 7 para la de matemáticas.

Los resultados de los cuadros 6 y 7 muestran que en términos brutos y en promedio los niños que asisten a establecimientos particulares subvencionados obtienen 17.8 puntos adicionales en la prueba SIMCE de castellano que los niños de los establecimientos municipales. Por su parte los niños que asisten a establecimientos particulares pagados obtienen 59.5 puntos más que los de establecimientos municipales. Los resultados para la prueba de matemáticas son similares.

Estas diferencias de puntaje se reducen en forma importante cuando se corrige por el índice de nivel socioeconómico de los estudiantes. Al incluir esta variable en la ecuación el diferencial en castellano disminuye a 12 puntos al comparar municipales y particulares subvencionados y disminuye aún más (de 59.5 a 22.9 puntos) al comparar municipales con particulares pagados. En matemáticas ocurre algo similar, el diferencial entre municipales y particulares pagados se reduce de 63 a 27.5 puntos. Esto significa que parte importante, si bien no todo, del diferencial de puntaje en la prueba SIMCE entre los niños establecimientos de distinta dependencia puede ser explicado por las diferencias de nivel socioeconómico.

Al agregar otras variables que caracterizan a los establecimientos y el sexo de los estudiantes los diferenciales entre los establecimientos municipales y los particulares subvencionados se mantienen prácticamente iguales. Por su parte, los diferenciales entre establecimientos municipales y particulares pagados se reducen sólo entre dos y cuatro puntos, en mayor medida en castellano que en matemáticas.

Los coeficientes estimados de las variables que caracterizan a los establecimientos nos permiten observar que los estudiantes que asisten a establecimientos científico-humanistas tienen mejores resultados que quienes asisten a establecimientos polivalentes y éstos a su vez mejores puntajes que los estudiantes de establecimientos técnico-profesionales. Esta situación se produce tanto en castellano como en matemáticas.

¹² Estos problemas se minimizan al estimar modelos en valor agregado, lo que como ya se explicó es infactible en Chile por falta de información adecuada.

¹³ El R^2 ajustado es alrededor de 26% en los modelos estimados a nivel de alumnos y 67% en los estimados a nivel de establecimientos. Los estudios internacionales realizados a nivel de alumnos tienen un R^2 ajustado que fluctúa entre 30% y 40%.

**Cuadro 6: Función de producción educacional a nivel de alumnos
(variable dependiente: prueba castellano SIMCE 2° Medio 1998)**

Variable	Modelo 1	Modelo 2	Modelo 3
Constante	238.87 (1001.34)**	248.54 (960.51)**	231.66 (116.80)**
Dummy P. Subvencionado	17.83 (50.90)**	12.19 (35.64)**	12.11 (24.78)**
Dummy P. Pagado	59.47 (108.41)**	22.87 (31.23)**	19.08 (9.68)**
Modalidad CH			14.94 (34.41)**
Modalidad Polivalente			3.61 (7.13)**
MECE 94			-14.89 (-7.98)**
MECE 95			-15.44 (-8.33)**
MECE 96			-7.27 (-4.02)**
MECE 97			-3.54 (-1.95)
MECE 98 y más			3.68 (1.83)
Fin. Compartido			5.99 E-04 (13.16)**
Jornada Completa			3.36 (7.04)**
Matrícula			7.61 E-03 (8.66)**
Matrícula al cuadrado			-1.44 E-06 (-5.41)**
Tasa alumno/prof.aula			-0.16 (-7.24)**
Exp. docentes			0.81 (21.26)**
Indice Socioeconómico alumno		22.13 (94.79)**	16.83 (66.23)**
Indice Socioeconómico alumno al cuadrado		-2.99 (-28.76)**	-2.24 (-20.21)**
Género alumno, mujer			5.77 (17.85)**
R ² ajustado	0.14	0.23	0.26
F	6069.8**	5482.3**	1350.5**
n	77796	74337	69402

Notas: Estadísticas t entre paréntesis, **variable significativa al 1%. Variables dummy de referencia: establecimientos municipales, modalidad TP, no participa en el programa MECE, establecimientos sin jornada completa, alumnos hombres.

Cuadro 7: Función de producción educacional a nivel de alumnos (variable dependiente: prueba matemáticas SIMCE 2° Medio 1998).

Variable	Modelo 1	Modelo 2	Modelo 3
Constante	238.33 (989.27)**	245.94 (925.95)**	233.28 (117.0)**
Dummy P. Subvencionado	16.62 (47.08)**	11.81 (33.75)**	11.67 (23.33)**
Dummy P. Pagado	62.99 (114.11)**	27.46 (36.71)**	24.99 (12.61)**
Modalidad CH			13.55 (30.50)**
Modalidad Polivalente			2.45 (4.74)**
MECE 94			-12.59 (-6.72)**
MECE 95			-14.52 (-7.81)**
MECE 96			-6.26 (-3.45)**
MECE 97			-0.97 (-0.54)
MECE 98 y más			6.52 (3.22)**
Fin. Compartido			5.77 E-04 (12.38)**
Jornada Completa			1.83 (3.73)**
Matrícula			8.64 E-03 (9.65)**
Matrícula al cuadrado			-1.95 E-06 (-7.22)**
Tasa alumno/prof.aula			-0.20 (-8.69)**
Exp. docentes			0.92 (23.72)**
Indice Socioeconómico alumno		18.68 (77.88)**	13.32 (51.06)**
Indice Socioeconómico alumno al cuadrado		-1.40 (-13.19)**	-0.57 (-5.00)**
Género alumno, mujer			-5.08 (-15.39)**
R ² ajustado	0.15	0.21	0.24
F	6604.2**	5009.9**	1218.7**
n	77678	74325	69440

Notas: Estadísticas t entre paréntesis, ** variable significativa al 1%. Variables dummy de referencia: establecimientos municipales, modalidad TP, no participa en el programa MECE, establecimientos sin jornada completa, alumnos hombres.

También es interesante destacar que aparentemente hay un tamaño óptimo de establecimiento, medido a través de su matrícula, esto se comprueba al observar que la matrícula tiene un efecto positivo sobre el puntaje en las pruebas, pero que esta relación no es lineal, sino cóncava es decir, luego de alcanzar un determinado tamaño, si el establecimiento aumenta su matrícula su rendimiento en las pruebas disminuye (o aumenta a una tasa mucho menor).

Asimismo, la variable índice socioeconómico tiene un efecto positivo sobre el desempeño escolar, pero su efecto no es lineal, llega un momento en que mayores niveles del índice socioeconómico incrementan muy poco o no incrementan el resultado en las pruebas SIMCE. Esto se comprueba al observar que la variable índice socioeconómico al cuadrado tiene un coeficiente negativo y estadísticamente significativo.

La tasa alumno-profesor de aula es negativa y estadísticamente significativa, lo que significa que a mayor número de estudiantes por profesor peor es el rendimiento de los estudiantes (nótese que este resultado se obtiene controlando por el tamaño del establecimiento).

A su vez, la experiencia de los docentes, medida en años de servicio, tiene un efecto positivo sobre los resultados de los estudiantes.

El hecho que el establecimiento tenga jornada completa también está positivamente correlacionado con el puntaje en las pruebas SIMCE de 2º medio, sin embargo, no es posible concluir acerca de la dirección de la causalidad de esta relación. Bien puede estar ocurriendo que los establecimientos con más y mejores recursos sean quienes han podido implementar la jornada completa; es decir, no podemos de estos resultados concluir que la jornada completa tiene un efecto positivo en los puntajes en el SIMCE.

Algo similar ocurre con la variable financiamiento compartido, la que tiene un efecto positivo y significativo en el puntaje obtenido en la prueba SIMCE. Esta variable puede asociarse a una mayor disponibilidad de recursos por parte de los padres y el establecimiento, así como también a una actitud de mayor compromiso de los padres con la educación de sus hijos. Es posible que la variable financiamiento compartido tenga más que ver con esto último – dado que el aumento en la disponibilidad de recursos no es tan relevante - lo que podría explicar su efecto positivo sobre los resultados escolares de los alumnos.

Un resultado interesante es que en castellano a las mujeres les va mejor que a los hombres, pero en matemáticas a los hombres les va mejor que a las mujeres, siendo las diferencias de puntaje entre ambos grupos muy similares.

El año de ingreso al MECE parece actuar como una variable proxy de las dificultades que enfrentan los establecimientos, es una variable con coeficiente negativo para aquellos establecimientos que ingresaron al MECE en los primeros años de este programa, precisamente porque eran quienes más lo necesitaban. Los establecimientos que han ingresado sólo recientemente al programa MECE en cambio no tienen coeficientes estadísticamente significativos, indicando que tienen resultados similares que aquellos que no están en el MECE, fundamentalmente los particulares pagados.

V. Conclusiones y comentarios finales

El análisis del rendimiento educacional, en base a la prueba SIMCE de 2° año de enseñanza media, confirma que los factores socioeconómicos son muy relevantes para explicar el resultado de los distintos tipos de establecimientos educacionales en las pruebas estandarizadas del SIMCE. Estos resultados se encuentran tanto en las estimaciones a nivel de establecimientos, como a nivel de estudiantes.

A nivel de establecimientos educacionales, es importante destacar también la significativa correlación que existe entre la dependencia del establecimiento y el nivel socioeconómico de los alumnos que asisten a él. Esta correlación es más marcada que la obtenida en estudios previos para enseñanza básica, lo que podría ser un indicador de una mayor segmentación social en los establecimientos de enseñanza media.

No obstante, estas no son las únicas variables que importan, una forma de mirar esto es a través del análisis de las diferencias entre el puntaje efectivo y el esperado, dadas las características de cada uno de los establecimientos educacionales. Lo que se observa es que hay dispersión al interior de cada tipo de educación, indicando que una vez que uno corrige por la variable socioeconómica hay otros elementos que también juegan un rol importante en el resultado obtenido por los establecimientos.

Asimismo, la diferencia entre el resultado efectivo y el esperado, dadas las características de los establecimientos, presenta una distribución distinta según la dependencia del establecimiento. Los establecimientos municipales, en promedio, tienen puntajes efectivos más bajos que los esperados en relación a los otros dos tipos de establecimientos. Por su parte, no hay diferencias estadísticamente significativas entre las distribuciones de las brechas de puntaje de los establecimientos particulares subvencionados y pagados.

A nivel de alumnos, además de las variables socioeconómicas importa el sexo de éstos, resultando diferenciales de signo opuesto, en castellano las mujeres tienen mejor rendimiento que los hombres y lo contrario sucede en matemáticas. Asimismo, importan determinadas características del establecimiento: modalidad, dependencia, tamaño y la tasa alumno-profesor. Así como las características de los profesores, que en este caso sólo hemos podido controlar a través de sus años de servicio.

Los resultados obtenidos para 2° Año de Enseñanza Media en 1998, indican que, a diferencia de lo que ocurre en 4° año de enseñanza básica en 1996, una vez que se comparan estudiantes de características similares (al menos en las variables que es posible incluir en el modelo) los alumnos que asisten a establecimientos municipales tienen menores resultados en el SIMCE que quienes asisten a establecimientos particulares subvencionados. Por su parte, los establecimientos particulares pagados, en promedio, tienen mayores puntajes que los particulares subvencionados.

Las razones de estas diferencias deberían explorarse más a fondo en el futuro. En primer lugar, es importante estudiar si estas diferencias se deben al nivel educacional (básica versus media) o si éstas reflejan mejoras de los establecimientos particulares

subvencionados en el tiempo. Esto podrá investigarse cuando se disponga de los nuevos resultados de la prueba SIMCE de 4° Básico de 1999.

A continuación planteamos algunas hipótesis que podrían explicar las diferencias de resultados entre 4° de Enseñanza Básica en 1996 y 2° de Enseñanza Media en 1998.

Una posible explicación es el menor tiempo transcurrido entre el comienzo del programa MECE-Media y la medición de logro del SIMCE 1998. Otra explicación posible es que en el caso de la enseñanza básica en 1996, aquellas escuelas municipales que tenían un mejor logro educativo que las particulares subvencionadas se encontraban localizadas en la zona rural, ocurriendo lo opuesto en la zona urbana; en el caso de la educación media la gran mayoría de los establecimientos educacionales se encuentran en la zona urbana. Asimismo, en la enseñanza media los padres tienden a preocuparse en mayor medida por el establecimiento al que asisten sus hijos, ya que piensan que su calidad afecta directamente las posibilidades futuras de éstos (por ej. el ingreso a la universidad). Los establecimientos de enseñanza media, a su vez, tienen mayor información para seleccionar a sus alumnos, ya que conocen su desempeño escolar en la enseñanza básica¹⁴.

La existencia de sesgo de selección entre los distintos tipos de establecimientos educacionales debe ser analizada en mayor profundidad. Este sesgo se basaría en el hecho que los padres no eligen aleatoriamente el establecimiento donde educan a sus hijos y los establecimientos tampoco eligen aleatoriamente a sus alumnos. Por lo tanto, sería interesante poder rehacer las estimaciones a nivel de alumnos corrigiendo por sesgo de selección, lo que nos permitiría determinar hasta qué punto los niños que asisten a un determinado tipo de establecimiento educacional tienen desventajas relativas con respecto al resto de los estudiantes¹⁵. El intento por realizar este ejercicio fue infructuoso, en gran medida por la falta de información necesaria para construir un modelo de elección de establecimientos que permita estimar adecuadamente la variable de selección. En la medida que el MINEDUC continúe mejorando sus bases de datos se podrá avanzar en este tipo de investigación¹⁶. Asimismo, se podrán estimar mejores funciones de producción educacionales en la medida que se cuente con información acerca de los recursos efectivos con que cuenta cada establecimiento educacional.

Adicionalmente, es necesario investigar las razones que explican el mejor desempeño relativo de los establecimientos particulares subvencionados en la enseñanza media. Además de lo mencionado previamente, es necesario investigar el efecto de las diferencias en gestión y administración de los recursos sobre el logro académico. Sin embargo, podemos adelantar, de acuerdo a nuestro análisis estadístico a nivel de establecimientos, que las diferencias de logro entre las dependencias no se explicarían por la existencia de liceos municipales de muy bajos puntajes, sino por la menor densidad de la cola derecha de

¹⁴ Los establecimientos particulares subvencionados pueden seleccionar a sus alumnos; sin embargo, sólo aquellos establecimientos municipales con exceso de demanda los pueden hacer (como es el caso, por ejemplo, del Instituto Nacional y del Liceo Carmela Carvajal).

¹⁵ Ver por ejemplo, Jimenez, Lockheed y Wattanawaha (1988), Glewwe y Jacoby (1993), Goldhaber (1996), Mizala, Romaguera y Reinaga (1999) quienes corrigen por este sesgo de selección.

¹⁶ Ya se ha avanzado significativamente al poder contar con información socioeconómica a nivel de alumnos.

la distribución, es decir, por la existencia de pocos liceos municipales con puntajes superiores a los esperado.

Este estudio muestra que uno de los desafíos fundamentales de la política educacional en los próximos años es reducir la brecha existente en el logro educativo entre los distintos tipos de establecimientos de enseñanza media. Es necesario lograr una mayor equidad, lo que implica que los aprendizajes deben mejorar más aceleradamente en los establecimientos que atienden a una población estudiantil de menor nivel socioeconómico. Asimismo, hay que evitar la segmentación que se tiende a producir al concentrarse los estudiantes con mayores dificultades económicas en determinados tipos de establecimientos. Es importante en este punto enfatizar que el cierre de la brecha en la calidad educativa no debe implicar un descuido de aquellos establecimientos con financiamiento público que tienen un elevado logro educativo; pues son precisamente estos liceos quienes juegan un rol fundamental en la movilidad social.

Referencias

- Aedo, C., "Organización industrial de la prestación de servicios sociales", *Documento de Trabajo* de la Red de Centros. BID, marzo 1997.
- Aedo C. y Larrañaga O., "Educación Privada versus Publica en Chile. Calidad y Sesgo de Selección", Mimeo 1994. Programa de Postgrado en Economía. ILADES/Georgetown University.
- Bellei, C., "Política hacia la educación media en los 90's. Un balance prov(m)isorio", MINEDUC, 1999.
- Berger M. y Toma E. "Variation in State Education Policies and Effects on Student Performance", *Journal of Policy and Management*, Vol. 13, N°3, 477-491, 1994.
- Benson, C.S., "The money we spend and what happens to it" en Gerwin, D. (ed) *The Employment of Teachers: Some Analytical Views*, Berkeley, McCutchan, 1974
- Borland, M. y R. Howsen, "Student achievement and the degree of market concentration in education" *Economics of Education Review*, vol 11, N° 1, 31-39, 1992.
- Bravo, D., Contreras, D. y C. Sanhueza, "Rendimiento educacional, desigualdad, y brecha de desempeño privado/público. Chile 1982-1997". *Documento de Trabajo* N°163, 1999.
- Card, D. y A. Krueger, "Labor market effects of school quality: Theory and evidence". In G. Burtless (ed) *Does Maney Matter? The Effect of School Resources on Student Achievement and Adult Success*. Washington D.C. Brookings Institution, 1996.
- Carnoy, M. y McEwan, "Public investments or private schools? A reconstruction of educational improvements in Chile", mimeo Stanford University, 1998.
- Chizmar, J. y Zak, T. "Modeling Multiple Outputs in Educational Production Functions", *American Economic Review*, mayo 1983.
- Carnegie Forum on Education and the Economy, *A nation prepared: Teachers for the 21st Century*. New York, Carnegie Corporation, 1986.
- Deller, S. y R. Rudnicki, "Production Efficiency in Elementary Education: The Case of Maine Public School", *Economics of Education Review*, Vol. 12. N°1, 1993.
- Ehrenberg, R. y D. Brewer, "Do school and teacher characteristics matter? Evidence from high school and beyond", *Economics of Education Review*, vol 13, N° 1, 1-17, 1994.
- Ehrenberg, R. y D. Brewer, "Did teachers' verbal ability matter in the 1960's? Coleman Revisited", *Economics of Education Review*, vol 14, N° 1, 1995

- Fuller, B. y P. Clarke, "Raising schools effects while ignoring culture. Local conditions and the influence of classroom, tools, rules and pedagogy" *Review of Educational Research*, vol 64, N° 1, 1994.
- Glewwe, P. y H. Jacoby, "Student achievement and schooling choice in low-income countries. Evidence from Ghana" *The Journal of Human Resources*, 843-864, 1993.
- Goldhaber, D., "Public and private high schools: Is school choice an answer to the productivity problem?" *Economics of Education Review*, vol. 15, N° 2, 93-109, 1996.
- Hanushek, E., "The economics of schooling: Production and efficiency in the public schools" *Journal of Economic Literature* 24, 1986.
- Hanushek, E., "Interpreting recent research on schooling in developing countries" *The World Bank Research Observer*, vol 10, N°2, 227-246, 1995
- Harbison R. y Hanushek, E, *Educational Performance of the Poor: Lesson from Rural Northeast Brazil*, World Bank Oxford University Press 1992.
- Hanushek, E. y Taylor L., "Alternative Assesments of the Performance of Schools. Measurement of State variations in Achievement", *The Journal of Human Resources*, Vol. 25, N°2, 1990.
- Jimenez, E., Lockheed, M. y N. Wattanawaha, "The relative efficiency of private and public schools: The case of Thailand" *The World Bank Economic Review*, vol 2, N° 2, 139-163, 1988.
- Krueger, A., "Experimental estimates of education production functions", *NBER Working Paper* N° 6051, 1997.
- Mancebo, M.E., "La calidad de la educación básica en Uruguay, aspectos metodológicos de la investigación de la CEPAL", *Revista de la CEPAL* 1993.
- Meyer, R. "Can schools be held accountable for good performance? A critique of common educational performance indicators" en E. Hoffman (ed) *Essays on the Economics of Education* W.E. Upjohn Institute for Employment Research, 1993.
- Mizala, A. y P. Romaguera, School performance and choice: The Chilean Experience, *The Journal of Human Resources*, vol 35 N°2, Spring 2000.
- Mizala, A. y P. Romaguera, "¿Cómo se comparan los resultados de la prueba SIMCE entre colegios públicos y privados", *Revista Perspectivas en Política, Economía y Gestión*, vol 2 N°1, 1998.
- Mizala, A., Romaguera, P. y D. Farren, "Eficiencia técnica de los establecimientos educacionales en Chile", *Documento de Trabajo* N°38, Centro de Economía Aplicada, Depto. de Ingeniería Industrial, Universidad de Chile, 1998.

- Mizala, A, Romaguera, P. y T. Reinaga, “Factores que determinan el desempeño escolar en Bolivia”, *Documento de Trabajo* N°61, Centro de Economía Aplicada, Depto. de Ingeniería Industrial, Universidad de Chile, 1999.
- Peterson, T.K., “Designing accountability to help reform” en C. Finn y T. Rebarber (ed)*Education Reform in the 90's*. Mac Millan Press,1992.
- Rodriguez, J. “School achievement and decentralization policy: the Chilean case” *Revista Análisis Económico*, vol 3, N° 1, 75-88, 1988.
- Sander, W., “Catholic grade schools and academic achievement” *The Journal of Human Resources*, vol 31, N°3, 540-548, 1995.
- Summers, A. y Wolfe, B. “ Do school made a difference?”, *American Economic Review*, septiembre 1977.
- Wetzel, J. O’Toole, D. y Millner, E. “A qualitative Response Model of Student Performance on a Standarized Test “, *Atlantic Economic Journal*, Vol. 19, 1991.

Anexo.

Cuadro A1. Correlaciones entre el índice socioeconómico, educación de los padres e ingreso familiar a nivel de establecimientos

	Índice socioeconómico	Educación de los padres
Educación de los padres	0.963** (1661)	
Ingreso familiar	0.963** (1661)	0.856** (1661)

Nota: ** La correlación es significativa al 1% (bilateral).

El número entre paréntesis corresponde al número de establecimientos.

**Cuadro A2. Función de producción educacional a nivel de establecimientos
(variable dependiente: prueba SIMCE castellano 2° Medio 1998)**

Variable	Coficiente	Estadístico t
Constante	250.17	95.86**
Modalidad CH	4.55	2.90**
Modalidad Polivalente	2.56	1.43
MECE 94	- 1.85	- 0.53
MECE 95	- 3.20	- 0.98
MECE 96	4.54	1.59
MECE 97	8.99	3.31**
MECE 98 y más	8.68	2.51*
Fin. Compartido	0.0002	1.60
Jornada completa	1.70	1.18
Matrícula	- 0.0002	- 0.06
Matrícula al cuadrado	- 3.64E-07	- 0.28
Tasa alumno/prof. aula	0.11	1.94
Exp. Docentes	0.24	2.02*
Indice Socioeconómico	33.14	25.20**
Indice Socioeconómico al cuadrado	- 5.79	-10.57**
Establecimientos de hombres	6.17	2.80**
Establecimientos de mujeres	14.31	8.38**
R ² ajustado	0.70	
F	204.80	
n	1495	

Nota: **variable significativa al 1%, * variable significativa al 5%.

Variables dummy de referencia: modalidad TP, no participa en el programa MECE, establecimientos sin jornada completa, establecimientos mixtos.

**Cuadro A3. Función de producción educacional a nivel de establecimientos
(variable dependiente: prueba SIMCE matemáticas 2° Medio 1998)**

Variable	Coficiente	Estadístico t
Constante	246.09	89.62**
Modalidad CH	1.48	0.90
Modalidad Polivalente	1.10	0.59
MECE 94	0.34	0.09
MECE 95	- 1.13	- 0.33
MECE 96	6.09	2.03*
MECE 97	10.90	3.81**
MECE 98 y más	9.76	2.68**
Fin. Compartido	0.00009	0.64
Jornada completa	2.43	1.60
Matrícula	0.0002	0.51
Matrícula al cuadrado	- 9.11E-07	- 0.65
Tasa alumno/prof. aula	0.12	2.05
Exp. Docentes	0.38	3.06**
Indice Socioeconómico	31.75	22.94**
Indice Socioeconómico al cuadrado	- 2.74	- 4.75**
Establecimientos de hombres	15.58	6.73**
Establecimientos de mujeres	7.73	4.30**
R ² ajustado	0.67	
F	175.17	
n	1495	

Nota: **variable significativa al 1%, * variable significativa al 5%.

Variables dummy de referencia: modalidad TP, no participa en el programa MECE, establecimientos sin jornada completa, establecimientos mixtos.

Cuadro A4. Correlaciones entre el índice socioeconómico, educación de los padres e ingreso familiar a nivel de alumnos

	Índice socioeconómico	Educación de los padres
Índice socioeconómico		
Educación de los padres	0.882** (74337)	
Ingreso familiar	0.882** (74337)	0.555** (74337)

Nota: ** La correlación es significativa al 1% (bilateral).
El número entre paréntesis corresponde al número de alumnos.

Cuadro A5. Estadísticas descriptivas de las variables a nivel de alumnos prueba SIMCE de castellano

Variable	Media	Desviación típica	N
Prueba de castellano	252.31	49.11	69402
Dummy P. Subvencionado	0.41	0.49	69402
Dummy P. Pagado	9.29 E-02	0.29	69402
Modalidad CH	0.46	0.50	69402
Modalidad Polivalente	0.18	0.38	69402
MECE 94	0.12	0.33	69402
MECE 95	0.16	0.37	69402
MECE 96	0.34	0.47	69402
MECE 97	0.25	0.43	69402
MECE 98 y más	3.1 E-02	0.17	69402
Fin. Compartido	1847.0	4164.1	69402
Jornada Completa	0.14	0.35	69402
Matrícula	924.23	667.21	69402
Matrícula al cuadrado	1299369.1	1953345.65	69402
Tasa alumno/prof.aula	24.06	8.84	69402
Exp. docentes	13.43	5.0	69402
Índice Socioeconómico alumno	-3.79 E-02	0.96	69402
Género alumno, mujer	0.52	0.50	69402

Cuadro A6. Estadísticas descriptivas de las variables a nivel de alumnos prueba SIMCE de matemáticas

Variable	Media	Desviación típica	N
Prueba de matemáticas	251.71	49.63	69440
Dummy P. Subvencionado	0.41	0.49	69440
Dummy P. Pagado	9.33 E-02	0.29	69440
Modalidad CH	0.46	0.50	69440
Modalidad Polivalente	0.18	0.39	69440
MECE 94	0.12	0.33	69440
MECE 95	0.16	0.37	69440
MECE 96	0.34	0.47	69440
MECE 97	0.24	0.43	69440
MECE 98 y más	3.08 E-02	0.17	69440
Fin. Compartido	1847.64	4171.94	69440
Jornada Completa	0.14	0.35	69440
Matrícula	923.57	668.05	69440
Matrícula al cuadrado	1299268	19611382.89	69440
Tasa alumno/prof.aula	24.12	8.89	69440
Exp. docentes	13.40	4.99	69440
Indice Socioeconómico alumno	-3.8 E-02	0.96	69440
Género alumno, mujer	0.52	0.50	69440