

DESEMPEÑO Y BRECHA EDUCATIVA EN CHILE: ¿EXISTE UN SESGO POR COBERTURA?*

Andrea Gutiérrez E.**
Ricardo D. Paredes***

I. INTRODUCCIÓN

El fuerte incremento de la cobertura educacional que han observado la mayoría de los países latinoamericanos y Chile en particular en las últimas tres décadas, es sin duda uno de los logros más relevantes de las políticas educativas recientes. En contraposición, el diagnóstico generalizado en estos países es que la calidad de la educación dista muchísimo de lo que debería ser y que existen grandes brechas en la calidad recibida según el ingreso, lo que sugiere una perpetuación de la inequitativa distribución del mismo. En el caso de Chile, este diagnóstico es avalado por una gran cantidad de estudios y existe consenso en que, además, las brechas de calidad, a pesar del importante gasto público en educación, que entre 1990 y el 2008 se incrementó en más de 300% en términos reales y que ha sido particularmente dirigido a los sectores más vulnerables, no se han reducido (véase Brunner y Elacqua, 2003).

Aunque la persistencia de la baja calidad de la educación y de las brechas según niveles de ingreso parece un fracaso de la política educacional, una hipótesis natural es que ambos aspectos están relacionados y, en concreto, que el aumento de la cobertura explica en gran medida el estancamiento del promedio educacional y la persistencia de la brecha según nivel de renta. El aumento de la cobertura, como cabe esperar, ha significado ir incorporando a jóvenes especialmente vulnerables, vulnerabilidad que no capturan en forma completa las mediciones tradicionales realizadas por las encuestas de hogares. En consecuencia, los promedios en los tests de rendimiento académico debieron ir reflejando cada vez más la media poblacional, a

partir de reflejos de medias muestrales sesgadas por sobrerrepresentación de la población privilegiada, que excluía a los más vulnerables.

El problema anterior se conoce como sesgo de selección y surge del uso de muestras no aleatorias. La magnitud del sesgo y su dirección dependen del tipo de población que se excluye del análisis. Por ejemplo, si los jóvenes que asisten a las escuelas, y que por ende son observados rindiendo las pruebas, tienen características no observables diferentes de los que no asisten (por ejemplo, motivación familiar al estudio) y ellas se relacionan (positivamente) con el rendimiento escolar, entonces las medias de rendimiento de las muestras observadas no reflejarán la media poblacional y la media de la muestra estará sesgada hacia arriba. Como la incorporación progresiva de más jóvenes a la educación debió ir reduciendo estos sesgos, dos hipótesis de interés surgen de la relación previa y son materia de estudio en este trabajo: i) la evolución media de los tests de rendimiento no describe el comportamiento de poblaciones similares, sino de una muestra que se ha ido haciendo cada vez más representativa y, por ende, el progreso efectivo ha sido superior al observado en los tests medios de rendimiento, y ii) la brecha en la calidad de la educación está particularmente afectada por la cola inferior de la distribución, que se ha ido incorporando en forma progresiva, con lo

* *Agradecemos al Ministerio de Educación por el acceso a los datos del Simce; a Mideplán por el acceso a la encuesta Casen; los comentarios y sugerencias de André Portela Souza, María Carolina Leme, Rómulo Chumacero, Valentina Paredes y dos árbitros anónimos de Economía Chilena. Este proyecto se desarrolló con el financiamiento de Fondecyt, proyecto 1095176, y del Centro de Estudios de Políticas y Prácticas en Educación, CIE01-CONICYT.*

** *Dirección de Presupuesto, Ministerio de Hacienda. E-mail: agutierrez@dipres.gob.cl*

*** *Escuela de Ingeniería, Pontificia Universidad Católica de Chile. E-mail: rparedes@ing.puc.cl*

que la evolución de la brecha medida está sesgada hacia arriba.

En este trabajo evaluamos la existencia de este sesgo de selección y medimos su impacto en términos del efecto en las medias de rendimiento educacional y en las brechas según ingreso. El trabajo se organiza en cuatro secciones, aparte de esta introducción. En la segunda sección se hace una breve descripción de la literatura y de la situación educacional en Chile. La tercera sección presenta la metodología y los datos utilizados, y la cuarta presenta los resultados. La quinta sección concluye.

II. ANTECEDENTES GENERALES

1. Cobertura

La tasa de cobertura educacional en Chile para enseñanza media, definida como el porcentaje de jóvenes en edad escolar que asisten a los establecimientos educacionales, está afectada por la obligatoriedad legal de matrícula desde el 2003. El gráfico 1 muestra la evolución de las tasas de cobertura en el período 1990-2009.

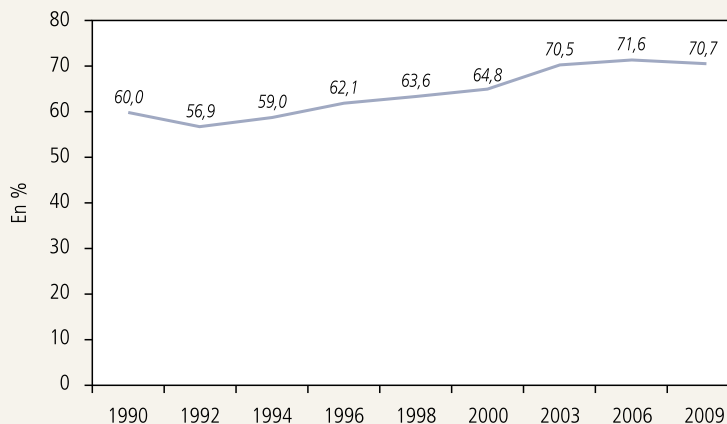
Como se puede apreciar, los niveles de asistencia a la enseñanza media han aumentado en el período reciente. Al comparar, por ejemplo, la cobertura del año 2000 con el último dato conocido, corres-

pondiente al 2009, se advierte un aumento de casi 6 puntos porcentuales. Entre los motivos que justifican este crecimiento, destacan el esfuerzo estatal por lograr, al igual que en los niveles básicos, el acceso universal de la población a la educación, y de este modo servir como una herramienta de movilidad social, con énfasis en los sectores más vulnerables.

El cuadro 1 muestra la evolución de la cobertura (o asistencia) por quintiles de ingreso. Como era de esperar, esta se ha incrementado especialmente en aquellos de menor ingreso. La diferencia en los

GRÁFICO 1

Asistencia Neta a la Educación Media Período 1990-2009^a



Fuentes: Casen, Mideplan.

a. Asistencia Neta a la Educación Media: número total de alumnos de 14 a 17 años que asisten a la educación media sobre la población de 14 a 17 años. Se excluye la educación especial.

CUADRO 1

Asistencia Neta a la Educación Media Período 1990-2009^a por Quintil de Ingreso Autónomo Per Cápita del Hogar

Año/quintil	Enseñanza media					Total
	I	II	III	IV	V	
1990	51.0	55.8	61.5	69.2	77.8	60.1
1998	49.9	61.5	67.3	73.3	83.1	63.6
2003	61.5	68.4	73.7	79.1	81.8	70.4
2006	62.9	71.0	74.6	78.2	82.1	71.6
2009	65.4	71.9	71.9	74.1	73.9	70.7
Dif. 2009-1990	14.4	16.1	10.4	4.9	-3.95	10.6

Fuente: Casen, Mideplan.

a. Asistencia Neta a la Educación Media: número total de alumnos de 14 a 17 años que asisten a la educación media sobre la población de 14 a 17 años. Se excluye la educación especial.

aumentos de cobertura por quintil sugiere que la población que se educa se ha ido tornando cada vez más representativa, lo que sugiere que los eventuales sesgos de muestra deberían haber ido cayendo en el tiempo.

Respecto de los factores que inciden en la asistencia escolar, sin duda el ingreso es un determinante relevante y, en general, hay un porcentaje importante para el cual la asistencia es una materia de decisión. Los resultados de la Encuesta de Caracterización Socioeconómica Nacional (Casen) 2009 indican, por ejemplo, que quienes están inhabilitados no superan el 3%, mientras que quienes han decidido no asistir por razones económicas superan el 30%. El 20% no asiste por razones de embarazo, y el resto no lo hace por razones que tienen sólo algún componente económico.

2. Aspectos Institucionales y Calidad de la Educación

La discusión sobre la calidad de la educación en Chile es antigua y, aunque relativamente poco documentada, dio origen a la reforma de los años ochenta (Cox, 2005). Hasta ese período, la educación estaba fuertemente centralizada, teniendo el Ministerio de Educación un rol muy relevante en la educación pública. Las negociaciones eran centralizadas con los profesores, y dicha centralización se consideraba una fuente de conflicto e inequidad regional. La reforma de 1980 buscó revertir esta situación que se reflejaba en una baja calidad de la educación, altas tasas de repetición y deserción, baja inversión y falta de incentivos.

Una de las principales herramientas de la reforma fue la introducción de incentivos a través de elementos de mercado, la focalización del gasto y, en particular, la privatización y descentralización. Se generaron así tres categorías de escuela: i) las escuelas municipales, administradas por las municipalidades con financiamiento proveniente de un subsidio basado en la asistencia de los alumnos y aportes de las municipalidades; ii) las escuelas privadas subvencionadas, financiadas por medio del mismo subsidio por asistencia por alumno; y iii) las escuelas privadas pagadas, que no reciben subvenciones del gobierno y operan con fondos de los apoderados.

Con el objeto de promover la competencia e inducir el mejoramiento de la calidad, en 1988 se implementó la prueba estandarizada Simce¹ (Sistema de Medición de la Calidad de la Educación), la cual es en la actualidad el principal indicador de evaluación de la calidad de las escuelas. En 1996 se incorporó el Sistema Nacional de Evaluación del Desempeño (SNED), que exige a las escuelas información sobre sus procesos y resultados educativos y establece incentivos a los profesores² y, en el año 2003, después de un largo proceso de negociación con el Colegio de Profesores, se inició en Chile la implementación del Sistema de Evaluación del Desempeño Profesional Docente, el cual se aplica a los profesores de establecimientos municipales y cuyo objetivo es mejorar la calidad de los profesores, insumo fundamental en el proceso de enseñanza. En el año 2006, después de una reforma curricular que cambió planes y programas y definió contenidos mínimos obligatorios, se finalizó la implementación del sistema de jornada escolar completa, que aumentó de 30 a 38 horas semanales el estudio en los cursos de enseñanza básica, y de 36 a 42 horas semanales en la enseñanza media.

Sobre las consecuencias de las reformas en los resultados académicos la literatura es abundante. La mayoría de los estudios sobre rendimiento en Chile han seguido modelos de amplia aceptación en la literatura (por ej., Coleman et al., 1966), que consideran factores sociodemográficos y de la función de producción de las escuelas. La mayor parte de ellos halla efectos significativos de esas variables, en particular de aquellas asociadas a la educación de los padres y del ingreso, variables que no son fáciles de modificar.³

La discusión inicial sobre los efectos de las reformas y, en particular, del sistema de *vouchers* por una parte

¹ Estas pruebas son rendidas por alumnos de cuarto y octavo básico, y segundo medio. Los cursos que las rinden lo hacen en forma alternada año tras año. Así, el 2004 la rindieron los alumnos de 8° básico, el 2005 los alumnos de 4° básico, y el 2006 los alumnos de 4° básico y II° Medio.

² Para una descripción, véase González (2000).

³ Véase, para distintas aplicaciones a Chile y sin ser exhaustivos, Mizala et al., 2002; Gallego, 2002; Hsieh y Urquiola, 2006; Paredes y Paredes, 2009; Anand et al., 2009; García y Paredes, 2010, y Drago y Paredes, 2011, con una revisión de trabajos).

ha tenido detractores directos como Carnoy (1997), y Carnoy y De Moura (1997), que sostienen que ellas no se reflejaron en incrementos de la calidad. Otros han centrado su cuestionamiento en el tema distributivo, como Hsieh y Urquiola (2006) que sostienen que, si bien hubo mejoras, estas no se distribuyeron uniformemente entre los distintos tipos de colegios, lo que creó mayores inequidades en el sistema educacional. De hecho, la discusión más reciente en Chile respecto del caso de la medición de desempeño, ha estado marcada por distintas visiones sobre la eventual selección que hacen las escuelas de los alumnos y las consecuencias sobre las estimaciones.

Más allá de lo anterior, los datos recientes obtenidos de pruebas estandarizadas no dan cuenta de avances significativos en el desempeño académico. Los resultados de la prueba Simce, por ejemplo, han estado virtualmente estancados durante los últimos diez años, con variaciones que no superan los 7 puntos en todos los niveles de enseñanza (menos de un décimo de una desviación estándar) y, por el lado de la dispersión por tipo de establecimiento y, en particular, respecto de la reducción de brechas por ingreso, tampoco se observan avances significativos. La convicción de que existe un déficit de calidad en relación con los recursos gastados en educación es también ratificada en relación con los resultados de otras pruebas (Beyer, 2001) y es evidente en la discusión sobre el deterioro progresivo de la matrícula de la educación municipalizada (Paredes y Pinto, 2009).

III. METODOLOGÍA Y DATOS

Siguiendo el grueso de la literatura, nuestra aproximación es estimar un modelo de rendimiento educacional del tipo:

$$R_i = \beta X_i + u_i \quad (1)$$

donde R_i es el rendimiento educacional del individuo i medido a través de una prueba estandarizada, y que es función de un conjunto de variables observables X y de un error asociado a no observables u .

El problema de estimar (1) sólo con individuos que rindieron la prueba, es decir, únicamente con aquellos que participan del sistema educacional, es

que es muy probable que el error esté correlacionado con X , lo que genera un sesgo. Este es un problema especialmente serio para propósitos de analizar la evolución del rendimiento en el tiempo, cuando tal exclusión varía, como de hecho ocurre en el caso de la enseñanza media.

En efecto, el problema de selección es que el rendimiento sólo será observado ($s_i = 1$) si el valor presente neto de educarse (VPN), es mayor que cero; de otra forma, $s_i = 0$. Esto es:

$$s_i \begin{cases} 1 & \text{si } VPN_i \geq 0 \\ 0 & \text{si } VPN_i < 0 \end{cases} \quad (2)$$

Como s depende de los beneficios y costos de educarse, incluyendo los no observables, podemos formular s_i como:

$$s_i = \gamma Z_i + e_i \quad (3)$$

Así, s_i es una variable continua latente que representa la propensión marginal a participar en el sistema educacional y que depende de un vector de variables que incluyen características personales y familiares (Z_i) y un término de error (e_i). Esto implica que el puntaje esperado para un alumno con características X , condicional a que participa, no depende sólo de las características incluidas en el vector X_i , sino también de las características personales y familiares incluidas en el vector Z_i .

Como el estimador MCO entrega la media no condicional a ser observada, las estimaciones por MCO estarán sesgadas de acuerdo con:

$$E(R_i | X_i) = \beta X_i + E(u_i | s_i > 0) \quad (4)$$

Siguiendo a Heckman (1979), para corregir este problema, estimamos una ecuación de participación mediante un modelo Probit. A partir de esta, computamos el inverso de la "razón de Mills" ($\hat{\lambda}$), que permite una corrección análoga a la inclusión de una variable omitida. Con ello, se obtiene (5), que provee estimadores insesgados de los parámetros.

$$E(R_i | s_i > 0) = \beta X_i + \beta_\lambda \hat{\lambda} + u_i \quad (5)$$

Estimamos (5) usando como variable dependiente los resultados de la prueba Simce de segundo medio para los años 2001, 2003, 2006 y 2008. No utilizamos datos previos debido a que recién se estandarizan los tests en 1998. Como especificación, seguimos una forma bastante estándar para Chile⁴, donde los regresores son el ingreso, la educación de los padres, el tipo de escuela, el sexo y si la escuela está en una zona urbana.

Adicional a la base Simce, utilizamos la información de la encuesta Casen realizada desde 1985 por el Mideplan, y cuyo principal objetivo es elaborar diagnósticos de la realidad socioeconómica. Como la prueba Simce sólo considera al grupo de individuos que efectivamente asiste al sistema educacional, recurrimos a la Casen —que es representativa de toda la población— para estimar los determinantes de la asistencia y, por ende, corregir por autoselección.

Por último, debido a que la Casen incluye datos para individuos de todas las edades, creamos bases homologables a las del Simce de segundo medio, acotando la muestra a personas con un rango de edad entre 14 y 19 años. Las estadísticas descriptivas de las bases de datos se presentan en los cuadros 2 y 3.

Como se puede apreciar en los cuadros mencionados, si bien existen diferencias en los valores medios de algunas variables de la Casen y el Simce, como por ejemplo la educación de los padres y el ingreso, estas resultan ser, en casi todos los casos, menores de $\frac{1}{4}$ de la desviación estándar correspondiente. Por otra parte, es importante mencionar que estas diferencias son esperables, por cuanto la Casen incluye información de individuos que están fuera y dentro del sistema educacional, mientras que el Simce sólo a aquellos que efectivamente se encuentran estudiando; por lo tanto, si se consideran las hipótesis planteadas en este trabajo, no es extraño que se observen estos contrastes entre ambas fuentes de información.

IV. RESULTADOS

En primer lugar estimamos la probabilidad de participación en el sistema educacional a través de un modelo Probit, considerando como regresores la escolaridad de cada uno de los padres (expresada en años), el ingreso del hogar, la zona de residencia

(Urbana=1), el número de escuelas por cada mil jóvenes en edad escolar,⁵ y las variables personales del alumno: edad, género (Hombre=1) y si este trabaja o no (Trabaja=1) obtenidas de la Casen. Los resultados de estas estimaciones en términos de sus coeficientes marginales se presentan en el cuadro 4, e indican que el modelo tiene un poder predictivo alto, el cual ha ido disminuyendo progresivamente en el tiempo. Esto es coherente con el hecho de que la educación media presenta menores tasas de participación y con que ella no fue obligatoria sino hasta el año 2003.

A partir de los resultados del cuadro 4, estimamos modelos estándares de rendimiento educacional, con y sin inclusión de la variable lambda proyectada ($\hat{\lambda}$).⁶ Estos modelos consideran como variables independientes la escolaridad de los padres, ingresos del hogar, ingresos al cuadrado, género (Hombre=1), zona geográfica (Urbana=1), *dummies* según la dependencia administrativa del colegio (particular subvencionado y particular pagado, con categoría base Municipal), y efectos aleatorios por comuna y por establecimiento, y se conocen como modelos jerárquicos lineales (HLM por su nombre en inglés).

Los cuadros 5 y 6 muestran los resultados para estas estimaciones (HLM y HLM con la Metodología de Heckman en dos etapas), para los tests de matemáticas y lenguaje. Ellas resultan coherentes con la mayor parte de los estudios realizados para Chile en términos de su signo y del impacto de las variables incluidas, y tampoco difieren de otras estimaciones que usan modelos HLM. Para nuestros propósitos, sin embargo, el resultado más importante es la sistemática significancia estadística reportada por la

⁴ Véase Mizala y Romaguera, 2000, y Drago y Paredes, 2011, para una revisión)

⁵ Esta variable se obtuvo del cruce entre el número de colegios de enseñanza media obtenido del Directorio de Establecimientos del Ministerio de Educación para cada año y comuna y el número de jóvenes con edades entre 14 y 19 años de la encuesta Casen correspondiente.

⁶ Esta variable lambda es obtenida aplicando la forma funcional de la ecuación de participación obtenida a partir de la Casen en las base Simce del año respectivo. Para evaluar las variables Edad y Trabaja, que no están disponibles en las bases Simce, asumimos que estas tomaban el valor promedio de quienes estudiaban según los datos de la Casen. La identificación es, entonces, por forma funcional y por las distintas variables en cada una de las ecuaciones.

CUADRO 2

Estadísticas Descriptivas según bases de datos de Encuesta Casen

Año	Quintil Estadístico	Estudia (%)	Escolaridad de la madre	Escolaridad del padre	Ingreso	Hombre		Trabaja	Colegios por cada 1000 jóvenes	
						Urbano (%)	Edad			
2000	I	Media D.E.	73.42 44.18	6.81 3.34	87,701 34,399	15.98 1.56	51.1 50.0	74.3 43.7	7.1 25.7	1.49 1.08
	V	Media D.E.	91.45 27.96	11.58 4.14	1,040,524 557,810	15.66 1.39	47.7 49.9	94.6 22.5	5.2 22.2	2.10 1.88
	Total	Media D.E.	81.11 39.14	8.60 4.08	401,071 421,843	15.87 1.53	51.9 50.0	84.5 36.2	7.4 26.1	1.64 1.36
2003	I	Media D.E.	76.54 42.37	7.02 3.38	100,708 35,510	16.41 1.26	52.2 50.0	76.1 42.7	6.1 23.9	2.16 1.37
	V	Media D.E.	90.97 28.66	11.74 3.92	1,144,956 643,629	16.14 1.13	54.4 49.8	94.1 23.5	6.2 24.1	2.87 2.25
	Total	Media D.E.	82.67 37.85	8.90 4.00	437,193 472,598	16.32 1.21	52.6 49.9	85.2 35.5	8.1 27.3	2.32 1.65
2006	I	Media D.E.	82.75 37.79	7.48 3.52	129,849 44,115	15.81 1.47	50.0 50.0	76.4 42.4	4.8 21.5	1.74 0.93
	V	Media D.E.	93.52 24.63	11.81 4.07	1,420,880 780,116	15.67 1.33	54.0 49.8	93.5 24.7	6.3 24.4	2.18 1.52
	Total	Media D.E.	85.99 34.71	9.05 3.99	547,660 579,250	15.81 1.43	51.9 50.0	86.3 34.4	6.6 24.8	1.86 1.16
2009	I	Media D.E.	79.05 40.69	7.86 3.60	150,504 66,507	16.43 1.21	48.7 50.0	78.4 41.1	4.3 20.3	2.52 1.53
	V	Media D.E.	86.86 33.78	11.47 4.03	1,452,357 626,337	16.26 1.16	52.5 49.9	91.8 27.4	6.9 25.4	2.96 2.22
	Total	Media D.E.	82.51 37.99	9.23 3.92	607,914 539,525	16.36 1.21	51.9 50.0	86.0 34.7	6.7 25.0	2.62 1.76

Fuente: Elaboración propia a partir de bases de datos de Encuesta Casen.

CUADRO 3

Estadísticas Descriptivas, según Bases de datos del Simce

Año	Quintil	Estadístico	Puntaje		Escolaridad				Ingreso	Hombre	Urbano	Municipal	Particular	
			matemáticas	lenguaje	Madre	Padre	subvencionado	Particular						
2001	I	Media	227.95	233.80	3.04	2.30	52,315	50.6%	90.6%	44.8%	30.1%	10.2%		
		D.E.	44.10	45.74	3.45	2.58	53,973	50.0%	29.2%	49.7%	45.9%	30.2%		
	V	Media	289.69	285.24	14.54	15.76	883,102	50.8%	98.8%	12.1%	36.8%	42.5%		
		D.E.	58.72	49.17	3.81	3.87	683,666	50.0%	11.1%	32.6%	48.2%	49.4%		
	Total	Media	247.25	251.91	9.21	9.43	283,266	49.7%	95.6%	31.5%	37.7%	15.6%		
		D.E.	53.23	50.20	4.96	5.17	431,860	50.0%	20.6%	46.4%	48.5%	36.2%		
2003	I	Media	223.01	234.57	4.40	2.63	112,827	47.3%	94.8%	20.5%	32.9%	6.2%		
		D.E.	50.27	45.63	3.70	2.88	49,147	49.9%	22.2%	40.4%	47.0%	24.2%		
	V	Media	292.31	287.16	13.22	14.14	943,921	51.0%	98.7%	4.0%	44.3%	36.4%		
		D.E.	59.97	46.80	4.97	5.02	661,462	50.0%	11.2%	19.7%	49.7%	48.1%		
	Total	Media	246.80	253.85	9.06	9.07	324,945	49.4%	97.2%	13.6%	40.8%	13.5%		
		D.E.	60.05	49.98	4.89	5.33	436,128	50.0%	16.6%	34.2%	49.1%	34.1%		
2006	I	Media	229.66	237.87	2.20	1.59	39,197	48.2%	91.7%	38.1%	40.2%	8.4%		
		D.E.	59.70	48.63	3.38	2.43	56,107	50.0%	27.6%	48.6%	49.0%	27.8%		
	V	Media	297.09	286.76	14.22	15.00	1,072,719	49.3%	98.5%	9.6%	50.0%	32.8%		
		D.E.	61.99	49.91	2.52	2.51	713,482	50.0%	12.3%	29.5%	50.0%	47.0%		
	Total	Media	253.00	254.98	9.48	9.66	345,839	48.2%	95.8%	28.9%	45.2%	12.4%		
		D.E.	64.92	51.80	4.92	5.06	494,308	50.0%	20.0%	45.3%	49.8%	33.0%		
2008	I	Media	222.12	232.89	6.95	6.27	124,191	46.6%	93.3%	45.5%	34.1%	6.5%		
		D.E.	52.19	44.31	2.98	2.61	49,558	49.9%	25.0%	49.8%	47.4%	24.7%		
	V	Media	298.71	290.14	14.58	15.35	1,139,709	50.3%	98.5%	7.5%	48.5%	37.7%		
		D.E.	60.06	48.73	2.71	2.65	513,478	50.0%	12.2%	26.4%	50.0%	48.5%		
	Total	Media	253.29	257.05	10.91	11.02	411,099	49.1%	96.7%	26.0%	47.8%	13.3%		
		D.E.	62.41	50.60	3.63	3.77	443,495	50.0%	17.9%	43.8%	50.0%	34.0%		

Fuente: Elaboración propia a partir de bases de datos Simce.

CUADRO 4

Efectos Marginales del Modelo de Participación Escolar

Variable	Segundo medio			
	2000	2003	2006	2009
Escolaridad de la madre	0.0085 (86.17)**	0.0098 (95.53)**	0.0079 (110.96)**	0.0104 (104.04)**
Escolaridad del padre	0.0079 (75.41)**	0.0074 (66.52)**	0.0059 (75.91)**	0.0052 (45.81)**
Ingreso	0.0000 (75.60)**	0.0000 (43.36)**	0.0000 (53.62)**	0.0000 (25.68)**
Edad	-0.0799 (359.93)**	-0.0836 (291.99)**	-0.0595 (327.22)**	-0.1032 (349.25)**
Hombre	0.0321 (53.10)**	0.0253 (38.92)**	0.0213 (45.92)**	0.0131 (19.61)**
Urbano	0.0362 (39.50)**	0.0288 (29.58)**	0.0032 (4.96)**	-0.0077 (8.50)**
Trabaja	-0.4009 (186.53)**	-0.3861 (178.91)**	-0.3411 (169.72)**	-0.2808 (126.83)**
Colegios por niño en edad escolar	0.0067 (24.86)**	0.0070 (28.11)**	0.0008 (3.86)**	0.0013 (6.29)**
Pseudo R ²	0.36	0.34	0.34	0.31
Nº de observaciones	1,190,824	956,802	1,340,020	988,192

Fuente: Elaboración propia a partir de bases de datos de Encuesta Casen.

Valor absoluto del estadístico Z entre paréntesis.

*Significativo al 5%; **Significativo al 1%.

variable $\hat{\lambda}$ en todos los modelos, la que como cabía esperar, toma valores mayores que cero, indicando la presencia de sesgo de selección positivo. Ello indica que la estimación sin corrección genera una sobrevaloración de los puntajes observados respecto de los que se obtendrían al considerar la población completa y presumiblemente, una subvaloración de la brecha según niveles de ingreso.⁷

El cuadro 7 muestra comparativamente los promedios del Simce, condicionados y no condicionados al problema de selección de la muestra. Como las correcciones por autoselección de alumnos en la enseñanza media son las adecuadas para realizar comparaciones entre los

distintos años, del cuadro 7 se deduce que, si bien el sesgo de selección significa sobrestimar el puntaje Simce, la mejora ha sido sobreestimada. Es decir, contrario a lo que se esperaría, el progreso efectivo del Simce, revelado por los promedios o las estimaciones condicionadas al sesgo de selección, no ha sido superior a lo observado a través de los promedios y posiblemente ha sido inferior.

⁷ Las mismas estimaciones fueron realizadas para cuarto y octavo básico. En ellas, la variable $\hat{\lambda}$ aparece como no significativa en algunos periodos, lo que además se acompaña de cambios de signo en sus coeficientes, lo que sugiere la inexistencia de sesgo de selección. Ello, como se señaló, era esperable dadas las altas tasas de cobertura reportadas en estos niveles.

CUADRO 5

Estimaciones HLM y HLM con Metodología de Heckman del Puntaje Simce en Matemáticas

Variable	Segundo medio							
	2001	2003	2006	2008	2008			
Constante	200.32 (85.26)**	187.19 (71.16)**	192.88 (61.37)**	183.13 (54.63)**	191.56 (61.88)**	153.71 (45.79)**	183.47 (61.75)**	153.28 (39.03)**
Escolaridad de la madre	0.46 (16.07)**	0.77 (19.06)**	0.56 (23.23)**	0.86 (19.83)**	0.68 (17.24)**	2.07 (31.50)**	1.02 (25.45)**	2.15 (20.54)**
Escolaridad del padre	0.30 (10.69)**	0.59 (15.29)**	0.31 (13.76)**	0.53 (15.16)**	0.35 (8.88)**	1.46 (25.34)**	0.69 (17.43)**	1.27 (20.04)**
Ingresos del hogar	1.4E-05 (16.13)**	1.7E-05 (18.48)**	2.0E-05 (23.47)**	2.1E-05 (24.11)**	2.5E-06 (2.75)**	4.9E-06 (5.37)**	6.6E-06 (6.33)**	8.2E-06 (7.87)**
Ingresos ²	-3.49E-12 (10.39)**	-4.51E-12 (12.94)**	-5.31E-12 (15.07)**	-5.43E-12 (15.41)**	-1.2E-13 (0.34)	-1.2E-12 (3.43)**	-1.29E-12 (2.40)*	-2.02E-12 (3.73)**
Hombre	11.08 (52.46)**	12.14 (52.21)**	10.11 (46.35)**	10.86 (45.96)**	12.07 (51.42)**	14.85 (57.78)**	9.44 (40.64)**	10.70 (41.78)**
Urbano	14.19 (6.42)**	16.19 (7.32)**	17.89 (5.81)**	18.81 (6.12)**	22.49 (7.49)**	22.48 (7.63)**	18.24 (6.30)**	17.34 (6.01)**
Particular subvencionado	22.39 (15.00)**	22.34 (15.03)**	25.27 (14.66)**	25.23 (14.67)**	26.98 (15.20)**	25.65 (14.73)**	27.66 (16.87)**	27.51 (16.85)**
Particular	52.50 (30.02)**	51.82 (29.72)**	56.59 (26.26)**	56.28 (26.16)**	66.40 (27.77)**	62.91 (26.79)**	63.81 (28.35)**	62.79 (28.00)**
$\hat{\lambda}$		25.93 (10.93)**		20.74 (8.24)**		86.98 (26.41)**		80.77 (11.68)**
Efectos aleatorios por comuna	Sí	Sí	Sí	Sí	Sí	Sí	Sí	Sí
Efectos aleatorios por establecimiento	Sí	Sí	Sí	Sí	Sí	Sí	Sí	Sí
N° de observaciones	195,135	195,135	219,542	219,542	223,576	223,576	194,455	194,455

Fuente: Elaboración propia a partir de bases de datos Simce y Encuesta Casen

Valor absoluto del estadístico Z entre paréntesis.

*Significativo al 5%; **Significativo al 1%

CUADRO 6

Estimación por HLM y HLM con Metodología de Heckman del Puntaje Simce en Lenguaje

Variable	Segundo medio							
	2001		2003		2006		2008	
Constante	217.05 (109.04)**	210.13 (89.71)**	213.61 (98.47)**	201.81 (83.34)**	215.48 (101.21)**	175.26 (72.88)**	207.44 (100.32)**	172.40 (55.93)**
Escolaridad de la madre	0.58 (19.84)**	0.75 (17.86)**	0.56 (25.28)**	0.92 (23.15)**	0.61 (17.69)**	2.08 (36.58)**	1.08 (30.34)**	2.39 (25.71)**
Escolaridad del padre	0.32 (11.07)**	0.48 (11.87)**	0.27 (12.83)**	0.54 (16.56)**	0.26 (7.59)**	1.44 (28.82)**	0.72 (20.29)**	1.39 (24.61)**
Ingresos del hogar	1.4E-05 (15.03)**	1.5E-05 (15.98)**	2.2E-05 (28.15)**	2.3E-05 (29.01)**	3.3E-06 (4.13)**	5.8E-06 (7.32)**	7.4E-06 (7.98)**	9.3E-06 (10.00)**
Ingresos ²	-4.43E-12 (12.75)**	-4.96E-12 (13.78)**	-6.47E-12 (19.99)**	-6.62E-12 (20.44)**	-7.28E-13 (2.41)*	-1.9E-12 (6.19)**	-2.06E-12 (4.31)**	-2.91E-12 (6.06)**
Hombre	-5.86 (26.85)**	-5.29 (22.04)**	-4.53 (22.56)**	-3.62 (16.67)**	-6.22 (30.58)**	-3.27 (14.69)**	-3.97 (19.21)**	-2.50 (10.98)**
Urbano	12.48 (6.58)**	13.54 (7.12)**	16.12 (7.59)**	17.22 (8.12)**	17.05 (8.27)**	17.07 (8.55)**	12.57 (6.27)**	11.52 (5.78)**
Particular subvencionado	21.00 (17.36)**	20.97 (17.38)**	19.09 (16.32)**	19.04 (16.34)**	20.72 (17.16)**	19.31 (16.54)**	18.59 (16.54)**	18.41 (16.52)**
Particular	40.49 (28.23)**	40.14 (28.01)**	36.66 (24.61)**	36.25 (24.39)**	45.47 (27.62)**	41.82 (26.18)**	40.89 (26.23)**	39.72 (25.64)**
$\hat{\lambda}$		13.64 (5.57)**		25.09 (10.84)**		92.36 (32.41)**		93.78 (15.23)**
Efectos aleatorios por comuna	Sí	Sí	Sí	Sí	Sí	Sí	Sí	Sí
Efectos aleatorios por establecimiento	Sí	Sí	Sí	Sí	Sí	Sí	Sí	Sí
Nº de observaciones	195,135	195,135	219,542	219,542	223,576	223,576	194,455	194,455

Fuente: Elaboración propia a partir de base de datos Simce y Encuesta Casen.

Valor absoluto del estadístico Z entre paréntesis.

*Significativo al 5%, **Significativo al 1%.

CUADRO 7

Puntajes Simce Promedios Estimados

Materia	Método	Segundo medio			
		2001	2003	2006	2008
Matemáticas	Sin corrección	246.1	246.0	250.1	248.5
	Con corrección	241.5	242.4	236.8	236.9
Lenguaje	Sin corrección	251.4	252.5	253.0	254.0
	Con corrección	248.9	248.1	239.0	240.6

Fuente: Elaboración propia a partir de estimaciones HLM y HLM con Heckman.

CUADRO 8

Brecha Educativa en Puntajes Simce

Materia	Método	Segundo medio			
		2001	2003	2006	2008
Matemáticas	Sin corrección	1.16	1.17	1.15	1.19
	Con corrección	1.20	1.20	1.33	1.27
Lenguaje	Sin corrección	1.14	1.13	1.11	1.15
	Con corrección	1.16	1.16	1.29	1.24

Fuente: Elaboración propia a partir de estimaciones HLM y HLM con Heckman.

1. Brechas de Resultados por Quintiles de Ingreso

Los resultados previos muestran la existencia de sesgo de selección para la enseñanza media, pero no dan cuenta ni de la magnitud de este sesgo en los distintos grupos de la población, ni de si ellos se traducen en variaciones de las brechas educativas en el tiempo. Lo esperable, según lo que revelan las cifras de cobertura, es una reducción del sesgo con el ingreso de la población, la que se traduciría en

un aumento efectivo superior al medido por simple comparación de medias, de las brechas de puntaje predicho entre personas de mayor y menor ingreso. Para testear la hipótesis de que efectivamente el cada vez menor sesgo de selección producido por el aumento de la cobertura generó un efecto estadístico sobre las brechas educativas, estimamos los puntajes medios para los distintos quintiles de ingreso del hogar con y sin corrección, y luego computamos el cociente, un indicador de brecha relativa entre el 20% más rico y el 20% más pobre de la población. El cuadro 8 presenta los resultados del indicador.

Tal como cabía esperar, los resultados indican que la magnitud del sesgo de medición, entendida como el diferencial entre el puntaje promedio predicho por HLM con la Metodología Heckman (con corrección) y el puntaje promedio predicho por HLM (sin corrección), es menor en el quintil de mayores ingresos. Lo anterior se traduce en que, para todas las materias y años, la brecha del puntaje corregido sea mayor con respecto a la observada (HLM), siendo especialmente pronunciada esta diferencia en la prueba de matemáticas.

Por último, para estimar la significancia de la evolución de la brecha en el tiempo, usamos *bootstrapping*. Los resultados se muestran en el cuadro 9, donde las celdas con asterisco muestran los casos en que la diferencia entre años en el indicador de brecha es estadísticamente significativa. Asimismo, se aprecia que ha existido aumento significativo en la inequidad en el puntaje promedio esperado, el cual resulta ser mayor estadística y económicamente al evaluar los años extremos de la muestra.

CUADRO 9

Evolución de la Brecha Educativa Corregida en Puntajes Simce

Año	Matemáticas				Lenguaje			
	2001	2003	2006	2008	2001	2003	2006	2008
2001			*	*			*	*
2003			*	*			*	*
2006				*				*
2008								

Fuente: Elaboración propia.

* Indica que existe diferencia estadísticamente significativa al 5% entre la brecha de los años correspondientes.

V. CONCLUSIONES

Una de las hipótesis más recurrentes pero inexplicadas en la literatura de educación en Chile —que ha habido un mejoramiento mayor en el rendimiento educacional, y que no está capturado por las medias de puntajes— surge de un antecedente que se repite en distintos países latinoamericanos, cual es el fuerte incremento de la cobertura educacional. Este incremento ha incorporado a personas relativamente excluidas no sólo del sistema educacional, sino del acceso a otros medios, incluyendo personas más pobres. En consecuencia, esta incorporación de una población menos privilegiada debería haber reducido el rendimiento medio de la población, incluso si la calidad provista por el sistema educativo no se hubiese deteriorado. En otros términos, los promedios simples estarían sujetos a un sesgo de selección de muestra, el que habría ido cayendo en el tiempo, pero que produciría dos efectos: i) una sobrestimación del rendimiento educacional, y ii) una subestimación de la caída de las brechas educativas según ingreso.

En este trabajo estimamos la existencia de tales sesgos y hallamos que, en efecto, hay una significativa sobrestimación de los puntajes cuando ellos se miden por el promedio en lugar de corregirlos por selección, y que alcanza fácilmente los 10 puntos (1/5 de la desviación estándar del Simce). Ello quiere decir que, si el sistema educativo chileno educara a una población típica y no a una relativamente privilegiada, los grados de aprendizaje serían muy inferiores a lo que son.

En segundo lugar, hallamos también que el sesgo —esto es, el diferencial entre el puntaje promedio predicho por Heckman y el puntaje promedio no condicionado— es menor en el grupo de mayor renta, lo que no sorprende debido a que entre quienes poseen mayor renta, la exclusión es sustancialmente menor. La consecuencia de lo anterior es que el efecto del ingreso sobre el aprendizaje o la brecha educativa asociada al ingreso es significativamente subestimado por los promedios en torno a los 10 puntos porcentuales para los quintiles de mayor y menor ingreso.

Por último, y contrario a lo esperado, no se aprecia que la mayor cobertura haya significado subestimar

la caída de las brechas reales en el tiempo. Sobre este hallazgo, no tenemos una explicación consistente.

REFERENCIAS

- Anand, P., A. Mizala y A. Repetto (2009). “Using School Scholarships to Estimate the Effect of Private Education on the Academic Achievement of Low-Income Students in Chile.” *Economics of Education Review* 28: 370–81.
- Beyer, H. (2001). “Falencias Institucionales en Educación: Reflexiones a Propósito de los Resultados del Timss.” *Estudios Públicos* N°82, otoño.
- Brunner, J.J. y G. Elacqua (2003). “Informe Capital Humano en Chile.” Escuela de Gobierno, Universidad Adolfo Ibáñez.
- Carnoy, M. (1997) “Is Privatization through Education Vouchers Really the Answer? A Comment on West.” *The World Bank Research Observer* 12(1): 105-16.
- Carnoy, M. y C. De Moura (1997). “¿Qué Rumbo Debe Tomar el Mejoramiento de la Educación en América Latina?” Documento presentado en *Seminario sobre reforma educativa*. Washington, D.C., Banco Interamericano de Desarrollo.
- Coleman, J.S. (1966). *Equality of Educational Opportunity*. U.S. Dept. of Health Education and Welfare Office of Education, Washington, D.C.
- Cox, C. (2005). *Políticas Educativas en el Cambio de Siglo*. Santiago, Chile: Editorial Universitaria.
- Drago, J.L. y R. Paredes (2011). “Meta-Análisis sobre Brecha de la Calidad de la Educación en Chile.” *Cepal Review*, por aparecer.
- Gallego, F. (2002). “Competencia y Resultados Educativos: Teoría y Evidencia Para Chile.” *Cuadernos. Economía* 39(118): 309–52.
- García, C. y R. Paredes (2010). “Reducing the Educational Gap in Chile: Good Results in Vulnerable Groups.” *Journal of Development Studies* 46(3).
- González, P. (2000). “Una Revisión de la Reforma del Sistema Escolar en Chile.” Documento de Trabajo N°92, Centro de Economía Aplicada, Universidad de Chile. Disponible en <http://econpapers.repec.org/paper/edjceauch/92.htm>
- Heckman, J. (1979). “Sample Selection Bias as a Specification Error.” *Econometrica* 47(1): 153–61.
- Hsieh, C.-T. y M. Urquiola (2006). “The Effects of Generalized School Choice on Achievement and

- Stratification: Evidence from Chile's Voucher Program.” *Journal of Public Economics* 90(8-9): 1477-503.
- Mizala, A. y P. Romaguera, R (2000). “Determinación de Factores Explicativos de los Resultados Escolares en Educación Media en Chile.” Serie Economía N° 85, Centro de Economía Aplicada, Departamento de Ingeniería Industrial, Facultad de Ciencias Físicas y Matemáticas, Universidad de Chile, agosto.
- Mizala, A., P. Romaguera y D. Farren (2002). “The Technical Efficiency of Schools in Chile.” *Applied Economics* 34(12): 1533-52.
- Paredes, R. y J. Pinto (2009). “¿El Fin de la Educación Pública en Chile?” *Estudios de Economía* 36(1): 47-66..
- Paredes, R. y V. Paredes (2009). “Educational Performance and Management under a Rigid Labor Regime.” *Cepal Review* 99: 117-29, diciembre.