

**MINISTERIO DE EDUCACIÓN**

**CONSTRUCCION DE GRUPOS  
COMPARABLES DE ESTABLECIMIENTOS  
SIMCE 8° BASICO 2000**

**Noviembre de 2001**

**DEPARTAMENTO DE ESTUDIOS Y ESTADÍSTICA  
DIVISIÓN DE PLANIFICACIÓN Y PRESUPUESTO**

## **CONSTRUCCION DE GRUPOS COMPARABLES DE ESTABLECIMIENTOS SIMCE 8° BASICO 2000**

### **I. Introducción**

Los resultados obtenidos en el SIMCE por los establecimientos educacionales están explicados por una multiplicidad de factores, los cuales pueden ser agrupados en variables internas a los establecimientos y externas. Las variables internas, como son la calidad docente, equipamiento e infraestructura y gestión del establecimiento, entre otras, son variables controlables y manejables por los establecimientos educacionales. Las variables externas al establecimiento como el nivel socioeconómico de los alumnos y el nivel educacional de los padres, entre otras, son variables o condiciones que afectan el rendimiento escolar del alumnado pero que el establecimiento educacional no controla, pues estas condiciones vienen dadas desde fuera, determinadas por el entorno político, cultural y/o socioeconómico en el que se desenvuelven los alumnos y sus familias. Por lo tanto, el resultado educativo obtenido por un establecimiento puede ser descompuesto y explicado por un conjunto de variables internas que dependen del establecimiento propiamente tal (lo que sucede “dentro” del establecimiento) y un conjunto de variables externas que no pueden ser afectadas por éste (lo que sucede “fuera” del establecimiento).

Cuando se desea comparar resultados educativos entre establecimientos educacionales es necesario tener en cuenta los aspectos mencionados anteriormente, de tal forma que las diferencias de resultados observadas entre establecimientos estén explicadas principalmente por lo que sucede “dentro” de los establecimientos y no por lo que sucede “fuera” de ellos.

Por lo tanto, si deseamos comparar resultados entre establecimientos y tener a la vez una mayor certeza que las diferencias de resultados que se observan entre éstos se están produciendo por lo que sucede “dentro del establecimiento” y no “fuera del establecimiento”, debemos comparar establecimientos que atiendan a alumnos de similares características socioeconómicas. Así, se puede decir con mayor certeza que, al observar diferencias de rendimiento dentro de un conjunto de establecimientos que atienden a segmentos de alumnos de similares características socioeconómicas, éstas se explican fundamentalmente porque existen diferencias respecto de lo que pasa “dentro” de los establecimientos. Por lo tanto, al interior de estos grupos de establecimientos resulta válido plantear que existen establecimientos que se desempeñan mejor que otros.

De lo anterior se desprende la necesidad y ventaja de agrupar establecimientos que enfrentan similares características externas previo a realizar comparaciones de resultados educativos. Una vez agrupados los establecimientos, se comparan los resultados educativos entre establecimientos pertenecientes a un mismo grupo.

Dentro de las características externas, según diversos estudios nacionales e internacionales, las más relevantes suelen ser aquellas que dan cuenta del nivel socioeconómico de los

alumnos: nivel educacional de los padres, nivel de ingreso familiar y grado de vulnerabilidad socioeconómica de los alumnos, entre las más importantes.

En concordancia con lo anterior, para la comparación de resultados del SIMCE 8° básico 2000, entre establecimientos de similares características en términos de la situación socioeconómica de los alumnos que atienden, se ha procedido a agrupar a los establecimientos utilizando las siguientes variables: nivel educacional del padre, nivel educacional de la madre, nivel de ingreso del hogar e índice de vulnerabilidad escolar. Las tres primeras variables provienen del cuestionario de padres o apoderados del SIMCE y la cuarta variable proviene de la JUNAEB<sup>1</sup>.

## **II. Metodología**

Para la construcción de los grupos de establecimientos comparables se utilizó la técnica estadística de análisis de conglomerados. Esta técnica genera grupos de establecimientos de tal forma que exista la mayor similitud socioeconómica al interior de cada grupo y la mayor diferencia, en términos de las variables socioeconómicas consideradas, entre éstos. A partir de lo anterior, se construyeron 5 grupos de establecimientos a nivel nacional.

## **III. Aplicación de la metodología: pasos, consideraciones y supuestos**

- De la base de datos del cuestionario de padres y apoderados del SIMCE 8° básico 2000 (en que los registros se encuentran a nivel de alumno) se tomaron las preguntas N° 5, N° 6 y N° 19, a partir de las cuales se generaron las variables años de escolaridad de la madre, años de escolaridad del padre e ingreso del hogar del alumno, respectivamente. Para las dos primeras variables se utilizó una tabla de conversión del nivel educacional (que se presenta a continuación) y el último curso aprobado en años de escolaridad cursados:

---

<sup>1</sup> Junta Nacional de Auxilio Escolar y Becas

**Tabla 1: Años de escolaridad asignados según nivel educacional y último curso aprobado.**

Nivel educacional	Ultimo curso aprobado							
	1	2	3	4	5	6	7	8
Básica	1 año	2 años	3 años	4 años	5 años	6 años	7 años	8 años
Media HC	9 años	10 años	11 años	12 años				
Media TP	9 años	10 años	11 años	12 años	12 años			
CFT	13 años	14 años	15 años	16 años	17 años			
IP	13 años	14 años	15 años	16 años	17 años			
Universidad	13 años	14 años	15 años	16 años	17 años	18 años	19 años	20 años
Magíster	18 años	19 años	20 años	21 años	22 años			
Doctorado	18 años	19 años	20 años	21 años	22 años	23 años	24 años	25 años

Para la generación de la variable de ingreso del hogar del alumno, se utilizó el punto medio de los intervalos de ingreso especificados en la pregunta N° 19 del cuestionario de padres y apoderados del SIMCE, según la siguiente tabla:

**Tabla 2: Ingreso imputado al hogar del alumno para cada tramo de ingreso.**

Tramo de Ingreso	Ingreso mensual imputado al hogar del alumno
Menos de \$ 100.000	\$ 50.000
De \$ 100.000 a \$ 200.000	\$ 150.000
De \$ 200.000 a \$ 300.000	\$ 250.000
De \$ 300.000 a \$ 400.000	\$ 350.000
De \$ 400.000 a \$ 500.000	\$ 450.000
De \$ 500.000 a \$ 600.000	\$ 550.000
De \$ 600.000 a \$ 800.000	\$ 700.000
De \$ 800.000 a \$ 1.000.000	\$ 900.000
De \$ 1.000.000 a \$ 1.200.000	\$ 1.100.000
De \$ 1.200.000 a \$ 1.400.000	\$ 1.300.000
De \$ 1.400.000 a \$ 1.600.000	\$ 1.500.000
De \$ 1.600.000 a \$ 1.800.000	\$ 1.700.000
Más de \$ 1.800.000	\$ 1.900.000

- A partir de la generación de las variables años de escolaridad de la madre, años de escolaridad del padre e ingreso del hogar del alumno se generaron tres nuevas variables a nivel de establecimiento: promedio de años de escolaridad de la madre, promedio de años de escolaridad del padre y promedio de ingreso mensual del hogar.

- Una vez generadas las variables promedio años de escolaridad de la madre, promedio años de escolaridad del padre y promedio ingreso mensual del hogar, se procedió a establecer un criterio de validez para los promedios calculados a nivel de establecimiento: los promedios calculados se consideraron como válidos si: 1º, el número de datos disponibles por variable es mayor o igual al 10% del número potencial de datos que se podría haber obtenido (Nº de cuestionarios distribuidos por establecimiento); y 2º, el número de datos disponibles para calcular los promedios es mayor o igual a 5. Si para algún establecimiento no se cumplían las 2 condiciones anteriores en alguno de los promedios calculados, el promedio calculado no se consideró (es decir, se trató como valor MISSING – dato faltante).
- Luego, a la base de datos con los promedios válidos de las variables promedio años de escolaridad de la madre, promedio años de escolaridad del padre y promedio ingreso mensual del hogar se le agregó el dato del IVE (Índice de Vulnerabilidad Escolar) de 1º básico 2000 y 1º medio 2000<sup>2</sup>. El IVE que se consideró finalmente fue el de 1º básico. Si no estaba disponible el IVE de 1º básico y estaba disponible el IVE de 1º medio, se consideró este último. Además se asignó un IVE igual a cero a todos los establecimientos particulares pagados.
- Una vez preparadas las variables promedio válidas por establecimiento (promedios válidos de las variables promedio años de escolaridad de la madre, promedio años de escolaridad del padre, promedio ingreso mensual del hogar e IVE), se procedió a estandarizar dichas variables, requisito previo para la aplicación del análisis de conglomerados. La estandarización se realiza con el propósito de que las variables estén cuantificadas en una misma escala de medición<sup>3</sup>.
- A través del programa estadístico SPSS y utilizando las 4 variables estandarizadas, se procedió a generar 5 grupos de establecimientos, que representan diferentes niveles socioeconómicos: bajo, medio-bajo, medio, medio-alto y alto, en base a los promedios de las 4 variables obtenidos a nivel de cada grupo y su comparación entre los grupos. Además, a partir de cuadros de resultados estadísticos entregados por el programa se chequeó que los 5 grupos generados fuesen significativamente distintos entre sí en relación a las variables consideradas.
- Todos los establecimientos con información para al menos una de las variables consideradas son asignados por el programa a alguno de los grupos.
- Finalmente se procedió a asignar aquellos establecimientos que no contaban con datos válidos para ninguna de las 4 variables consideradas (no pueden ser asignados

---

<sup>2</sup> Estos datos del IVE corresponden a los datos de vulnerabilidad recogidos el año 2000 para asignar las raciones del año 2001.

<sup>3</sup> Muchas medidas de distancia son sensibles a variaciones de escala o magnitud entre variables. Así, por ejemplo, si quisiéramos agrupar establecimientos por variables como la escolaridad promedio de los padres, ingreso promedio del hogar e índice de vulnerabilidad escolar, la mayor dispersión provendría de la variable ingresos y ocultaría las diferencias de escolaridad de los padres y vulnerabilidad escolar. La forma de proceder, es estandarizar las variables, es decir, restarles a cada una su promedio y dividir las por su desviación estándar.

por el programa por lo cual quedan omitidos de los grupos generados). Para clasificar a estos establecimientos en alguno de los grupos generados se utilizó en primera instancia un criterio de comuna-dependencia que consiste en chequear en qué grupo quedaron clasificados la mayoría de los establecimientos de la comuna y dependencia correspondiente al establecimiento no clasificado y asignarlo a dicho grupo. En caso de empate (que existiese el mismo número de establecimientos de la comuna y dependencia del establecimiento no clasificado en dos o más grupos) o en caso que no existiesen más establecimientos en la comuna con la dependencia del establecimiento no clasificado, se utilizó en segunda instancia el criterio de comuna que consiste en chequear en qué grupo quedaron clasificados la mayoría de los establecimientos de la comuna correspondiente al establecimiento no clasificado y asignarlo a dicho grupo.

#### **IV. Resultados finales**

Los grupos socioeconómicos construidos fueron designados por una letra: grupo A, grupo B, grupo C, grupo D y grupo E. El grupo A considera establecimientos que atienden a niños con un menor nivel educativo de los padres, un menor nivel de ingreso familiar y una mayor vulnerabilidad. El grupo E considera a establecimientos que atienden niños cuyos padres tienen un mayor nivel educativo, un mayor nivel de ingreso y una menor vulnerabilidad. Los grupos B, C y D presentan una situación intermedia de las características consideradas.

Para la comparación de resultados SIMCE 8° básico 2000 y 1997, se supuso que la conformación de los grupos socioeconómicos se mantuvo estable, es decir, que los establecimientos que quedaron clasificados en un determinado grupo, a partir de la clasificación de establecimientos obtenida con la información del año 2000, pertenecían a ese mismo grupo en el año 1997.

A continuación se presentan los cuadros con los resultados de las agrupaciones obtenidas. Se presenta:

- la distribución resultante de los establecimientos por grupo socioeconómico y dependencia,
- la distribución de la matrícula en educación regular total (matrícula inicial año 2000) de los establecimientos correspondientes por grupo socioeconómico y dependencia,
- la distribución de los alumnos que rindieron el SIMCE 8° básico 2000 por grupo socioeconómico,
- el promedio y la desviación estándar resultante para cada una de las 4 variables consideradas por grupo socioeconómico.

**CUADRO 1: NUMERO DE ESTABLECIMIENTOS POR GRUPO Y DEPENDENCIA**  
(en cursiva % fila y en negrilla % columna)

Grupo	Dependencia			Total
	Municipal	Particular Subvencionado	Particular Pagado	
A	967	200	0	1.167
(bajo)	82,9	17,1	0,0	100,0
	<b>33,4</b>	<b>12,5</b>	<b>0,0</b>	<b>23,1</b>
B	1.261	278	0	1.539
(medio-bajo)	81,9	18,1	0,0	100,0
	<b>43,6</b>	<b>17,3</b>	<b>0,0</b>	<b>30,5</b>
C	572	656	0	1.228
(medio)	46,6	53,4	0,0	100,0
	<b>19,8</b>	<b>40,9</b>	<b>0,0</b>	<b>24,3</b>
D	95	453	105	653
(medio-alto)	14,6	69,4	16,1	100,0
	<b>3,3</b>	<b>28,2</b>	<b>19,0</b>	<b>12,9</b>
E	0	18	448	466
(alto)	0,0	3,9	96,1	100,0
	<b>0,0</b>	<b>1,1</b>	<b>81,0</b>	<b>9,2</b>
Total	2.895	1.605	553	5.053
	57,3	31,8	10,9	100,0
	<b>100,0</b>	<b>100,0</b>	<b>100,0</b>	<b>100,0</b>

**CUADRO 2: MATRICULA 2000 EN EDUCACION REGULAR POR GRUPO Y DEPENDENCIA**  
(en cursiva % fila y en negrilla % columna)

Grupo	Dependencia			Total
	Municipal	Particular Subvencionado	Particular Pagado	
A	168.998	34.925	0	203.923
(bajo)	82,9	17,1	0,0	100,0
	<b>12,3</b>	<b>3,8</b>	<b>0,0</b>	<b>8,0</b>
B	651.283	124.875	0	776.158
(medio-bajo)	83,9	16,1	0,0	100,0
	<b>47,3</b>	<b>13,6</b>	<b>0,0</b>	<b>30,6</b>
C	455.385	419.875	0	875.260
(medio)	52,0	48,0	0,0	100,0
	<b>33,1</b>	<b>45,9</b>	<b>0,0</b>	<b>34,5</b>
D	100.515	326.656	33.319	460.490
(medio-alto)	21,8	70,9	7,2	100,0
	<b>7,3</b>	<b>35,7</b>	<b>13,5</b>	<b>18,1</b>
E	0	9.286	214.244	223.530
(alto)	0,0	4,2	95,8	100,0
	<b>0,0</b>	<b>1,0</b>	<b>86,5</b>	<b>8,8</b>
Total	1.376.181	915.617	247.563	2.539.361
	54,2	36,1	9,7	100,0
	<b>100,0</b>	<b>100,0</b>	<b>100,0</b>	<b>100,0</b>

**CUADRO 3: Nº DE ALUMNOS QUE RINDIERON SIMCE 8º BASICO 2000 POR GRUPO Y DEPENDENCIA**  
(en cursiva % fila y en negrilla % columna)

Grupo	Dependencia			Total
	Municipal	Particular Subvencionado	Particular Pagado	
A	19.149	3.903	0	23.052
(bajo)	<i>83,1</i>	<i>16,9</i>	<i>0,0</i>	<i>100,0</i>
	<b>13,3</b>	<b>4,5</b>	<b>0,0</b>	<b>9,2</b>
B	67.860	11.535	0	79.395
(medio-bajo)	<i>85,5</i>	<i>14,5</i>	<i>0,0</i>	<i>100,0</i>
	<b>47,1</b>	<b>13,3</b>	<b>0,0</b>	<b>31,6</b>
C	47.773	40.874	0	88.647
(medio)	<i>53,9</i>	<i>46,1</i>	<i>0,0</i>	<i>100,0</i>
	<b>33,2</b>	<b>47,0</b>	<b>0,0</b>	<b>35,2</b>
D	9.182	29.881	2.857	41.920
(medio-alto)	<i>21,9</i>	<i>71,3</i>	<i>6,8</i>	<i>100,0</i>
	<b>6,4</b>	<b>34,3</b>	<b>13,9</b>	<b>16,7</b>
E	0	863	17.741	18.604
(alto)	<i>0,0</i>	<i>4,6</i>	<i>95,4</i>	<i>100,0</i>
	<b>0,0</b>	<b>1,0</b>	<b>86,1</b>	<b>7,4</b>
Total	143.964	87.056	20.598	251.618
	<i>57,2</i>	<i>34,6</i>	<i>8,2</i>	<i>100,0</i>
	<b>100,0</b>	<b>100,0</b>	<b>100,0</b>	<b>100,0</b>

**CUADRO 4: PROMEDIO Y DESVIACION ESTANDAR DE LAS VARIABLES UTILIZADAS PARA GENERAR LOS GRUPOS**

Grupo		Promedio años escolaridad de la madre	Promedio años escolaridad del padre	Promedio ingreso del hogar (\$ 2000)	Promedio Índice de Vulnerabilidad Educativa (IVE - JUNAEB) (*)
A (bajo)	Promedio	6	7	92.334	70
	<i>Desviación Estándar</i>	<i>1</i>	<i>1</i>	<i>41.814</i>	<i>10</i>
B (medio-bajo)	Promedio	8	8	129.644	44
	<i>Desviación Estándar</i>	<i>1</i>	<i>1</i>	<i>35.876</i>	<i>9</i>
C (medio)	Promedio	10	11	198.329	23
	<i>Desviación Estándar</i>	<i>1</i>	<i>1</i>	<i>49.521</i>	<i>8</i>
D (medio-alto)	Promedio	13	13	409.106	7
	<i>Desviación Estándar</i>	<i>1</i>	<i>1</i>	<i>143.905</i>	<i>7</i>
E (alto)	Promedio	16	17	1.163.343	0
	<i>Desviación Estándar</i>	<i>1,1</i>	<i>1,1</i>	<i>298.241</i>	<i>0</i>

(\*) El valor del Índice de Vulnerabilidad Escolar refleja el porcentaje de alumnos en situación de vulnerabilidad para un determinado establecimiento.