



*Al servicio  
de las personas  
y las naciones*

Programa de las Naciones Unidas para el Desarrollo - Chile  
Área de Reducción de la Pobreza y la Desigualdad

## **CLASIFICACION SOCIOECONOMICA DE ESTABLECIMIENTOS PARA LEY DE ASEGURAMIENTO.\***

Oswaldo Larrañaga<sup>†</sup>  
Gustavo Cabezas<sup>‡</sup>

Documento de Trabajo  
Julio 2011

**Palabras clave:** vulnerabilidad, ficha de protección social, educación.

---

\* Las opiniones expresadas en este texto son de exclusiva responsabilidad del o los autores y no comprometen al Programa de las Naciones Unidas para el Desarrollo (PNUD).

<sup>†</sup> PNUD.

<sup>‡</sup> PNUD.

## CLASIFICACION SOCIOECONOMICA DE ESTABLECIMIENTOS PARA LEY DE ASEGURAMIENTO<sup>1</sup>

La condición socioeconómica de los hogares es un determinante principal de los resultados educacionales en el país, en particular en aquellos medidos a través de pruebas de conocimientos como son el Simce, Pisa, Timss y PSU. La variable socioeconómica se asocia a los recursos materiales y culturales que cuentan los hogares y que influyen decisivamente sobre el desarrollo infantil y la capacidad de aprendizaje de los niños. En ocasiones interesa distinguir entre aquella parte de los resultados de la educación que son atribuibles a la condición socioeconómica de los hogares de aquellos que se asocian a la calidad “intrínseca” de los establecimientos o a las acciones de la política pública. Un ejemplo es la publicación que realiza el Ministerio de Educación de los resultados del Simce al interior de grupos socioeconómicos. Otro caso es la asignación de recursos públicos a los establecimientos de mejor desempeño al interior de categorías socioeconómicas afines (Sned, subvención preferencial). También están las investigaciones académicas en educación que incorporan un control estadístico de la variable socioeconómica para aislar su efecto en la estimación de funciones de producción, efectividad de las escuelas y otros estudios relacionados.

Asimismo, la recientemente promulgada ley de aseguramiento de la calidad de la educación requiere que los establecimientos educacionales sean clasificados según niveles de logro en estándares educacionales, teniendo en cuenta la condición socioeconómica (vulnerabilidad) de los estudiantes en cada establecimiento. Específicamente, la referida ley establece en su artículo 17, inciso tercero:

“Para llevar a cabo esta ordenación, la Agencia deberá considerar los resultados de aprendizaje de todas las áreas evaluadas censalmente en las mediciones nacionales, la distribución de los resultados de los alumnos en relación con los estándares de aprendizaje y el grado de cumplimiento de los otros indicadores de calidad educativa. También deberá considerar las características de los alumnos del establecimiento educacional, incluida, entre otras, su vulnerabilidad, y, cuando proceda, indicadores de progreso o de valor agregado. Con todo, gradualmente, la ordenación de los establecimientos propenderá a ser realizada de manera independiente de las características socioeconómicas de los alumnos y alumnas, en la medida que el sistema corrija las diferencias atribuibles a dichas características en su desempeño.”

El objetivo del presente documento es presentar una metodología para medir el nivel socioeconómico de los establecimientos educacionales para su posterior utilización en la referida ordenación mandatada por la Ley de Aseguramiento de la Calidad. Al documento no le corresponde establecer cómo se utilizará la medida en la ordenación de las escuelas, sino que propone una medida que posee un conjunto de propiedades deseables para ser usada a tal

---

<sup>1</sup> Preparado por Osvaldo Larrañaga y Gustavo Cabezas del PNUD.

propósito. El indicador propuesto puede ser usado en contextos adicionales a la ley de aseguramiento de la calidad, puesto que sus ventajas analíticas tienen validez general.

La medida propuesta es el porcentaje de alumnos vulnerables por establecimientos, considerando como vulnerables a quienes poseen un puntaje en el Ficha de Protección Social por debajo de un determinado umbral. La vulnerabilidad puede entenderse como un riesgo elevado de caer en pobreza y se relaciona con los recursos y necesidades económicas que poseen los hogares. La Ficha de Protección Social es el instrumento que utiliza la política pública en el país para identificar los potenciales beneficiarios de los programas sociales focalizados, y consiste en una evaluación de los activos y necesidades económicas de los hogares.

La metodología requiere evaluar socioeconómicamente a cada alumno para clasificarlo como vulnerable o no, para luego agregar esta información a nivel de cada establecimiento.<sup>2</sup> La Ficha de Protección Social cubre en la actualidad a un 76% de la población escolar, lo cual permite cubrir razonablemente bien al universo de la población vulnerables en el sistema educativo. Ello si se considera que los hogares que no tienen Ficha pueden ser clasificados como no vulnerables, habida cuenta que la evaluación que se realiza a través del citado instrumento es requerida para postular a la mayor parte de los beneficios sociales de la política pública.

De tal manera, es posible calcular el porcentaje de alumnos vulnerables por establecimiento educacionales a partir de la información contenida en la Ficha de Protección Social. El indicador representa una medida más estable de la condición socioeconómica de los establecimientos que la entregada por la metodología de los grupos Simce en actual uso. Es también más flexible en sus aplicaciones, más fácil en su comunicación y más simple en su elaboración. Finalmente, es consistente con la evaluación socioeconómica de los hogares que realizan las políticas públicas para identificar a los beneficiarios de programas sociales focalizados.

El documento se organiza en seis secciones. En la primera sección se presentan las mediciones usadas actualmente en el país para medir la condición socioeconómica de estudiantes y establecimientos. En la segunda sección se presenta la medida propuesta y se argumenta sus ventajas respecto de las alternativas. En la sección tercera se desarrolla un análisis numérico de la distribución de establecimientos educacionales que resulta de la aplicación de la medida propuesta, usando a tal efecto una base de datos especialmente preparada por Mideplan. La cuarta sección analiza la distribución descrita a nivel de las dependencias educacionales y del tamaño de los establecimientos. La quinta sección realiza un análisis comparado en materia de umbrales alternativos para definir vulnerabilidad. Finalmente, en la sección sexta se detalla el funcionamiento de la Ficha de Protección Social y se adelantan los futuros desarrollos en esta materia.

---

<sup>2</sup> Los estudiantes también pueden clasificarse por grados de vulnerabilidad.

## 1.- La mediciones del nivel socioeconómico en actual uso

Las medidas en actual uso se pueden dividir en dos grupos, según operen a nivel de establecimientos educacionales o de alumnos individuales.

A nivel de los establecimientos educacionales se utiliza la metodología denominada Grupos Simce. Esta consiste en una clasificación de los establecimientos en cinco grupos socioeconómicos afines, construidos en base a las variables de escolaridad de los padres e ingreso familiar que se reportan en los cuestionarios de las pruebas Simce que deben llenar los padres o tutores de los alumnos. La metodología de los “grupos Simce” se basa en técnicas de clusters o conglomerados, que operan sobre la base de formar grupos afines de establecimientos según la cercanía estadística que presenten en las variables señaladas.

La clasificación de los Grupos Simce tiene tres usos principales:

- Comparación de resultados de las pruebas Simce a nivel de promedios para cada grupo socioeconómico. Así, las publicaciones del Mineduc comparan los resultados del Simce por dependencia educacional, al interior de los grupos socioeconómicos Simce.
- Subvención preferencial. Las escuelas son clasificadas en categorías según sus resultados al interior de los grupos socioeconómicos Simce; el nivel de clasificación resultante (autónoma, emergente y en recuperación) determina las condiciones que las escuelas deben cumplir para acceder a los recursos provistos por la subvención preferencial.
- SNED. Las escuelas acceden a los incentivos económicos del SNED a partir de los resultados que logran al interior de grupos homogéneos que se forman por región, zona, nivel de enseñanza y los referidos grupos socioeconómico Simce.

Por otra parte, hay un conjunto de mediciones de la condición socioeconómica de los *alumnos*, que tienen como común denominador estar basadas directa o indirectamente en la Ficha de Protección Social

En primer lugar, las escuelas adscritas al régimen de subvención preferencial reciben una transferencia adicional por cada alumno que califique en la categoría de prioritario. Esta categoría considera a los estudiantes pertenecientes al 40% de hogares de menores recursos, cuya identificación se realiza a través del puntaje que recibe su hogar en la Ficha de Protección Social (FPS). Para los alumnos cuyos hogares no tienen ficha de protección social (una minoría en la población de interés) se usan criterios alternativos como la pertenencia del hogar al sistema de protección social Chile Solidario, la escolaridad de la madre o el grupo Fonasa de pertenencia.

Segundo, la Junaeb utiliza el Sistema Nacional de Asignación por Equidad (Sinae) para seleccionar a los beneficiarios del Programa de Alimentación Escolar y demás beneficios de salud y útiles escolares. El Sinae califica a los estudiantes en cuatro categorías o prioridades; pertenecen a la

primera prioridad quienes estén participando en Chile Solidario (que a su vez han sido identificados por la Ficha de Protección Social), aquellos adscritos a la red del SENAME y quienes tienen Ficha de Protección Social con puntaje por debajo del umbral de indigencia (percentil 5) o que califiquen como indigente de acuerdo a mediciones alternativas. Califican en segunda y tercera prioridad los estudiantes cuyo hogar tenga un puntaje en la Ficha de Protección Social entre 4.213 y 8.500 puntos (percentil 5 a 20). Si la familia no tiene Ficha se usa el criterio Fonasa/escolaridad-madre referido. Para diferenciar entre segunda y tercera prioridad se usa un indicador de rendimiento académico, priorizándose en la segunda categoría a aquellos estudiantes que muestran un rendimiento más bajo.

En tercer lugar, en el año 2006 se decretó que los establecimientos subvencionados debían incorporar por lo menos a un 15% de alumnos vulnerables, a efectos de acceder a la subvención regular (o demostrar que no había postulantes suficientes en esa categoría). La calificación de vulnerabilidad se realiza a través de un promedio ponderado de 90% según puntaje en el ficha CAS (actualmente, Ficha de Protección Social) y 10% según pobreza regional. Los umbrales para calificar como vulnerables son específicos para cada región.

De esta manera, los Grupos Simce que operan a nivel de establecimientos se basan en la información de los cuestionarios para padres de la prueba Simce, mientras que las medidas a nivel de estudiantes están por lo general basadas en la información provista en la Ficha de Protección Social. Esta división de funciones para la información es solo circunstancial, puesto que ambas bases de datos pueden ser usadas para caracterizar socioeconómicamente tanto a estudiantes como establecimientos. De hecho, la medida que se propone en este documento es una clasificación de los establecimientos en base a la información de la Ficha de Protección Social. Asimismo, las investigaciones sobre los determinantes de los resultados individuales del Simce utilizan como variable socioeconómica aquellas que provienen de los cuestionarios Simce.

## **2.- El porcentaje de estudiantes vulnerables como medida del nivel socioeconómico de los establecimientos**

La propuesta del documento es caracterizar socioeconómicamente a los establecimientos educacionales a través del porcentaje de alumnos vulnerables. Para tal efecto se debe usar la información provista por la Ficha de Protección Social, que es una evaluación socioeconómica que se realiza a los hogares para identificar los beneficiarios de los programas sociales focalizados, y que asigna a cada hogar encuestado un puntaje que representa su nivel socioeconómico.<sup>3</sup>

La vulnerabilidad económica está asociada al riesgo que un hogar pueda experimentar situaciones de pobreza debido a que cuenta con una insuficiente dotación de activos socioeconómicos en relación a sus necesidades. La práctica en el país califica como vulnerables a los hogares que están

---

<sup>3</sup> En la sección 6 se provee una descripción más detallada de este instrumento.

por debajo de un determinado umbral de puntaje en la Ficha de Protección Social (tal como la pobreza se mide en términos de los hogares cuyo ingreso está por debajo de la línea de pobreza).

Actualmente, un 76% de los estudiantes de educación básica y media del país vive en hogares que han sido evaluados a través de la Ficha de Protección Social.<sup>4</sup> En caso que el hogar del estudiante no dispusiera de Ficha de Protección Social se le puede clasificar como no vulnerable, bajo el razonable supuesto que si no tiene Ficha es porque su hogar cuenta con recursos económicos suficientes y no necesita postular a los beneficios sociales del Estado. De esta manera, es posible clasificar a cada estudiante del sistema educativo como “vulnerable” o “no vulnerable”. Al interior del grupo de “vulnerables” se pueden distinguir niveles de intensidad de la condición, como se detalla más adelante.

El indicador propuesto resulta simplemente de calcular el porcentaje de estudiantes vulnerables en cada establecimiento, un cálculo trivial una vez que se ha clasificado a cada uno de los ellos.

Hay varias razones que sustentan la elección del indicador propuesto. En primer lugar, hay un argumento a favor de basar la medida en la información de la Ficha de Protección Social en vez de los cuestionarios Simce. Ello puesto que la Ficha provee información para el conjunto de los alumnos del establecimiento, mientras que los indicadores basados en los cuestionarios de Simce quedan referidos a los cursos que rinden la prueba. Esta característica resta representatividad a los indicadores derivados de la prueba Simce a la vez que los hace inestables en el tiempo, puesto que los factores aleatorios adquieren mayor importancia en muestras pequeñas. En cambio, un indicador que se base en el conjunto de los estudiantes del establecimiento resulta ser más representativo y estable, atributos deseables para la mayor parte de los usos que se puede entregar a la medición.<sup>5</sup>

Segundo, una vez elegida la FPS como base de datos se requiere justificar porque se trabaja con el porcentaje de alumnos vulnerables en vez de otras medidas posibles, como el promedio de la FPS a nivel de establecimiento, la mediana de la FPS, etc u otro. La ventaja del indicador propuesto es que puede ser calculado para todos los estudiantes del establecimiento, aún cuando no todos ellos tengan la evaluación de la Ficha, bajo el razonable supuesto que los hogares sin Ficha son hogares no vulnerables. En cambio, medidas como el promedio del puntaje de la FPS requieren que todos los estudiante del etablecimiento tengan un puntaje en la FPS.

Tercero, la medida es versátil en sus aplicaciones. El porcentaje de alumnos vulnerables puede ser usado en numerosos contextos, adaptandose a los objetivos del caso. Matemáticamente, se trata de una función que asocia establecimientos y valores numéricos en la escala 0 a 100. Esta variable

---

<sup>4</sup> En el Anexo A se presenta la distribución del puntaje de la Ficha en la población de estudiantes. La distribución tiene un punto modal en su límite superior, puesto que la escala de puntajes está truncada en ese punto (puesto que la información de los puntajes más altos no es relevante a los usos de la Ficha).

<sup>5</sup> En el Anexo B se analiza en mayor detalle la inestabilidad asociada a los Grupos Simce.

puede ser usada en procedimientos estadísticos como regresiones, sea como variable de control para aislar el efecto del factor socioeconómico, como variable dependiente de interés o como variable dependiente. Otro tipo de aplicación es la conformación de agrupaciones de establecimientos de similar nivel socioeconómico, a través de deciles, quintiles o similares.

Cuarto, la medida propuesta no le significa al Ministerio gastos en su elaboración, puesto que los datos ya han sido recolectados y es trivial realizar el cálculo por establecimiento, cuando se ha puesto en conjunto los datos de la Ficha de Protección Social y del Registro de Estudiantes de Chile (Rech). Así, no requiere ser externalizada a terceros ni actualizada en el tiempo, más allá de ir recalculando el porcentaje de alumnos prioritarios por establecimientos a medida que se cuenta con nueva información. Se trata también de una medida de muy fácil comunicación, en contraste a aquellas basadas en modelos estadísticos como regresiones o conglomerados.

Quinto, Mideplan está realizando un serio esfuerzo para que la Ficha de Protección Social esté desprovista de problemas de subdeclaración y relacionados en el futuro cercano. A tal efecto se están rediseñando tanto los procesos de recolección de la información en terreno como el modelo de cálculo del puntaje. En cambio, hay un riesgo que la información asociada al Simce pueda ser manipulada debido a que hay incentivos para ello, puesto que las escuelas tienen mayor probabilidad de acceder a más recursos u obtener mejores trato regulatorios si buenos resultados educativos respecto de otras escuelas en su misma categoría socioeconómica. Por tanto, es de su conveniencia competir en categorías socioeconómicas más bajas, y podrían “señalar” a los padres que la subdeclaración de información irá en su beneficio en la forma de mayores recursos para la educación de sus hijos.

Por último, la medida propuesta está basada en el mismo instrumento que se utiliza en la mayor parte de las políticas públicas que necesitan clasificar socioeconómicamente a la población. Ello contribuye a una mayor consistencia de la política pública y a una mejor comunicación con la comunidad. Tiene también la ventaja de economías de escala, puesto que el costo por usuario de generar y actualizar la información decrece con el volumen de uso.

### **3.- Distribución de establecimiento según el porcentaje de alumnos vulnerables**

Esta sección contiene un análisis cuantitativo de la medida propuesta. Específicamente, se presenta la distribución de establecimientos educacionales según el porcentaje de alumnos vulnerables, considerando escenarios alternativos para la definición de vulnerabilidad. La base de datos que se utiliza para el análisis numérico fue preparada especialmente por Mideplan y resulta de poner en conjunto los micro-datos de los registros administrativos de la Ficha de Protección Social y del Registro de Estudiantes de Chile del Ministerio de Educación. Ello se hace posible a través del registro único nacional (RUN), que es un identificador a nivel individual y que es común a ambas bases de datos.

La base de datos resultante contiene la información del puntaje de la Ficha de Protección Social para cada estudiante cuyo hogar haya sido sujeto de esta evaluación socioeconómica. La cobertura asciende a un 76% del total de estudiantes en enseñanza básica y media en el año 2010, porcentaje que fluctúa entre cerca del 8% para los establecimientos particular pagados y un rango de 85% a 90% para la educación municipal (Cuadro 1). La elevada cobertura de la Ficha de Protección Social en la población estudiantil es consistente con el alcance nacional del instrumento, que se estima en dos terceras partes de los hogares del país. La cobertura a nivel de los estudiantes es algo mayor que en el conjunto de la población, puesto que los hogares que tienen Ficha son los más vulnerables y en ellos reside un número relativamente elevado de niños y jóvenes.

**Cuadro 1: Alumnos con Ficha de Protección Social como porcentaje de la matrícula, 2010**

	matrícula	% con FPS
Municipio corporación	430.895	85,0
Municipio Daem	862.418	90,5
Particular subvencionado	1.491.942	74,8
Particular pagado	219.102	7,7
Corporación TP	54.631	86,1
Total	3.058.988	76,0

Fuente: Base de datos conjunta, FPS y RECH

A diferencia de la medida de pobreza que usa como umbral la línea de pobreza, la medición de la vulnerabilidad no cuenta con un criterio operativo bien definido en términos de un puntaje de la Ficha de Protección Social bajo el cual un hogar califique como vulnerable. Por ello, se trabaja con cuatro alternativas para luego recomendar una de ellas en base a los resultados obtenidos. Los puntajes potenciales son aquellos que usan los principales programas sociales focalizados para definir su población objetivo:

- 4.213 puntos, que es el puntaje utilizado para seleccionar a los beneficiarios de los programas sociales que se dirigen a la población más pobre, como es el caso de Chile Solidario.
- 8.500 puntos, que es usado para seleccionar a los beneficiarios de los programas orientados al 20% de población de menores recursos. Entre tales iniciativas se cuentan el Fondo Solidario para Vivienda I y los programas asistenciales de la Junaeb.



- 11.734 puntos, umbral que utilizan programas orientados a sectores de ingresos bajos y medio-bajos, incluyendo el Fondo Solidario de Vivienda II; Chile Crece Contigo; Subsidio Unico Familiar; Becas Nuevo Milenio y otros más
- Un promedio ponderado que entrega mayor peso a los estudiantes con menor puntaje de la FPS. Específicamente, se divide en deciles la población de estudiantes con menos de 11.734 en la FPS y se pondera por un factor de 10 a quienes están en el decil más pobre, por un factor de 9 al penúltimo decil y así sucesivamente hasta ponderar por un factor de 1 al decil más alto. El puntaje final se escala de modo de hacerlo comparable con las ordenaciones previas. Se usa un puntaje de 11.734 como límite superior puesto que la gran mayoría de los programas sociales focalizados utilizan umbrales igual o inferior a ese valor.

El análisis de mayor interés a efectos del objetivo del documento es la distribución de los establecimientos educacionales según el porcentaje de alumnos vulnerables. Para cada uno de los umbrales considerados se presente la distribución resultante en términos de percentiles; esto es, se ordena a los establecimientos según el porcentaje de alumnos vulnerable (de menor a mayor) y se reporta esta variable para determinados percentiles de la distribución.

El Cuadro 2 presenta un resumen de los resultados obtenidos, reportante el porcentaje de alumnos vulnerables a nivel del conjunto de los establecimientos para diferentes percentiles de la distribución. De esta manera, cuando se usa como umbral de vulnerabilidad 4.213 puntos en la FPS resulta que el establecimiento situado en el percentil 10 tiene un 7,4% de estudiantes vulnerables, el establecimiento en el percentil 25 tiene un 16,7% de vulnerabilidad, etc. <sup>6</sup>

Como se espera, un umbral más alto (exigente) para definir la condición de vulnerabilidad tiene asociados mayores porcentajes de estudiantes vulnerables para cada percentil. Así, cuando se usa un umbral de 11.734 puntos resulta que el establecimiento situado en el percentil 50 tiene un 84,2% de estudiantes vulnerables, comparado con solo un 25,9% del mismo percentil cuando se usa un umbral de 4.500 puntos. Lo mismo ocurre si se observan los promedios nacionales del porcentaje de estudiantes vulnerables: desde un 28,8% con el umbral más bajo a un 76,1% con el umbral más elevado. Por su parte, el umbral que se construye como un promedio ponderado tiene asociado un porcentaje de alumnos vulnerables en un nivel intermedio, lo que también es un resultado esperado

---

<sup>6</sup> Hay que notar que el porcentaje de alumnos vulnerable es bastante elevado respecto del porcentaje de población que teóricamente debiera estar por debajo de los respectivos umbrales. Ello sucede por la práctica de subdeclarar la condición socioeconómica, que sesga hacia abajo los puntajes de la FPS.

**Cuadro 2: Distribución del porcentaje de alumnos vulnerables**

	Puntaje FPS de umbral			
	4.213	8.500	11.734	Promedio ponderado
Percentil 10	7,4	23,8	38,5	19,1
Percentil 25	16,7	45,0	67,9	34,7
Percentil 50	25,9	63,3	84,2	47,0
Percentil 75	37,5	78,5	93,1	57,1
Percentil 90	53,6	91,7	100,0	67,5
Promedio	28,8	60,2	76,1	45,0

Fuente: Base de datos conjunta, FPS y RECH

La distribución completa de los establecimientos según el porcentaje de alumnos vulnerables se muestra en los gráficos 1 a 4, para los respectivos umbrales de vulnerabilidad. La inspección visual de las distribuciones otorga una clave para seleccionar el umbral que resulta más apropiado a los fines del ejercicio. El argumento tiene carácter funcional y se refiere a la capacidad que tiene la medida para discriminar en un sentido estadístico entre establecimientos. A tal respecto, una medida que entregue una distribución muy plana de establecimientos resulta poco funcional, porque no realiza distinciones suficientemente informativas entre establecimientos a los fines requeridos. Asimismo, una distribución que agrupe a la mayor parte de los establecimientos en áreas específicas de la distribución tampoco resulta atractiva cuando el propósito es poder distinguir entre establecimientos.

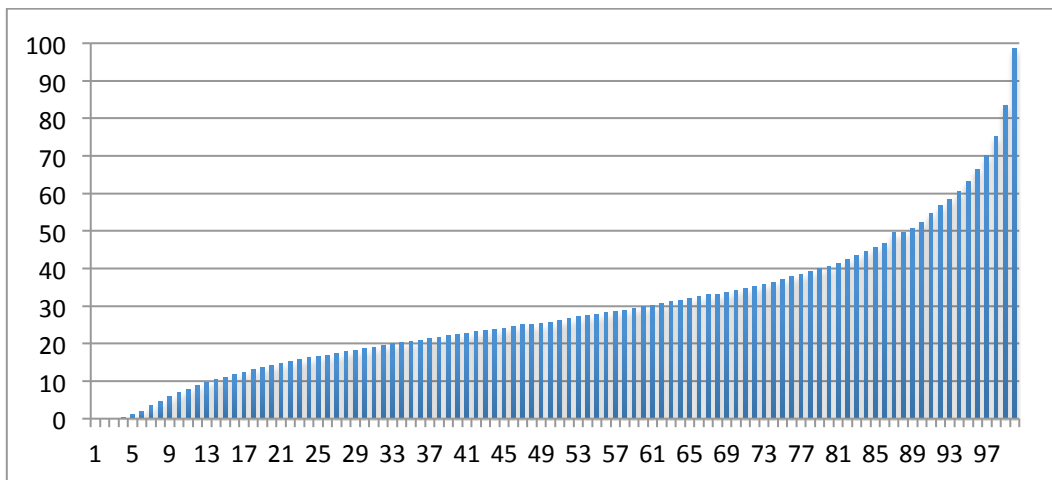
De tal manera, resulta que el umbral del promedio ponderado es el más informativo, puesto que da lugar a una distribución con mayor varianza y que ocupa la mayor parte del rango de la medida (entre 0% y 80% de alumnos vulnerables), sin presentar concentraciones de establecimientos en áreas específicas de la distribución. En cambio, el umbral de 4.213 puntos sitúa a la gran mayoría de los establecimientos en un rango más acotado (por debajo del 60% de vulnerabilidad) y entrega un mismo valor (0%) a los cinco primeros percentiles. De manera similar, los umbrales de 8.500 y 11.734 puntos concentran a la mayor parte de los establecimientos en la parte superior de la distribución, incluyendo a muchos que comparten el valor máximo del rango (100%).

La alternativa del promedio ponderado tiene también una ventaja conceptual sobre las otras. Esta reside en su carácter “continuo” en el sentido que ordena a los alumnos en deciles según puntaje en la FPS y les otorga un peso decreciente de modo que quienes están en los deciles de más bajos puntajes se consideran “más vulnerables” que aquellos en los deciles superiores. En cambio, las otras medidas tienen carácter “discreto” en el sentido que cada estudiante se le asocia un valor 0 o 1 según esté por arriba o debajo del respectivo umbral, lo que puede ser algo arbitrario para quienes están cerca del umbral. Además, el hecho que no existe una definición única y

consensuada del umbral de vulnerabilidad es un argumento adicional a favor de una medida “continua”, que representan una alternativa más ecléctica de medición que los umbrales discretos.

**Gráfico 1:**

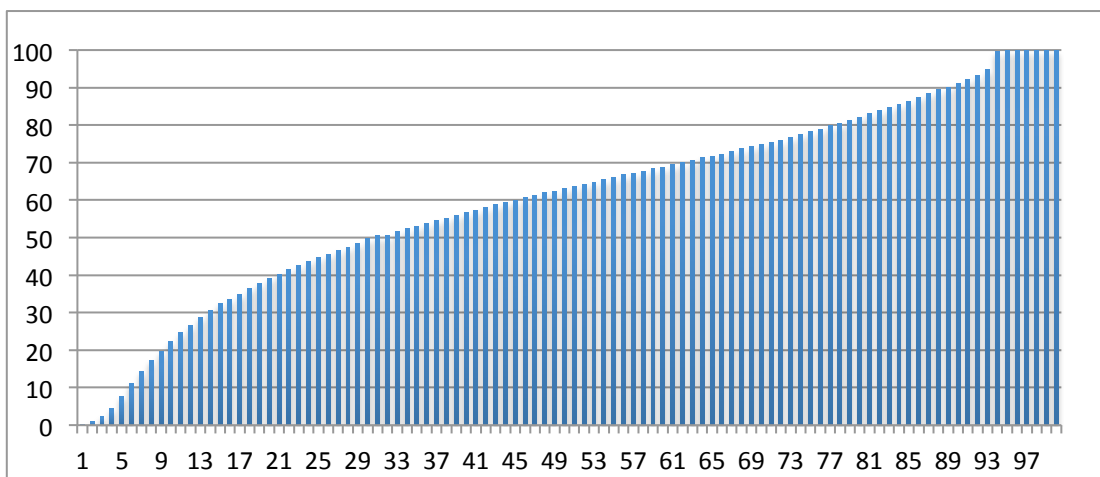
**Distribución de establecimientos según % de alumnos vulnerables. Umbral 4.213 puntos**



Fuente: Base de datos conjunta, FPS y RECH

**Gráfico 2:**

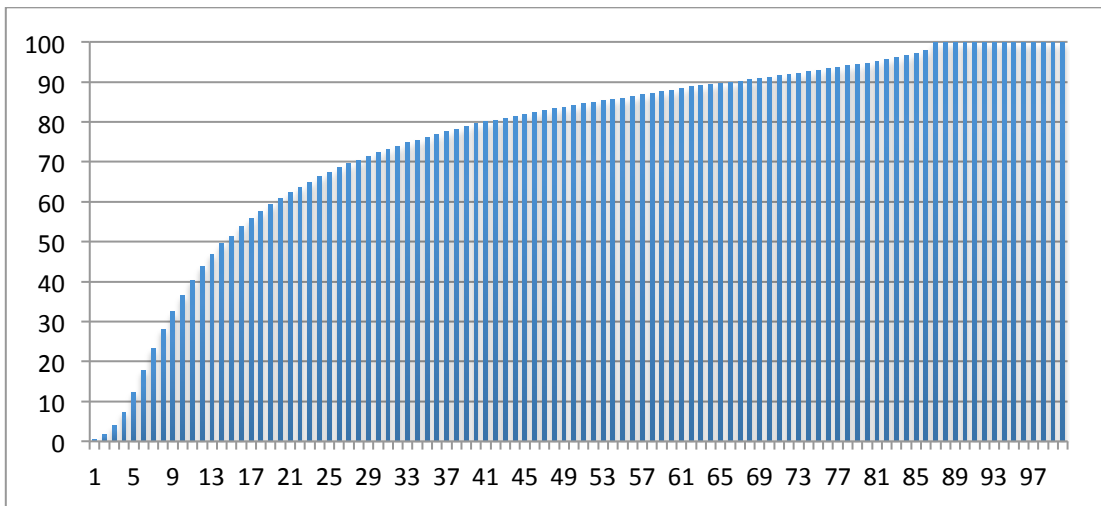
**Distribución de establecimientos según % de alumnos vulnerables. Umbral 8.500 puntos**



Fuente: Base de datos conjunta, FPS y RECH

**Gráfico 3:**

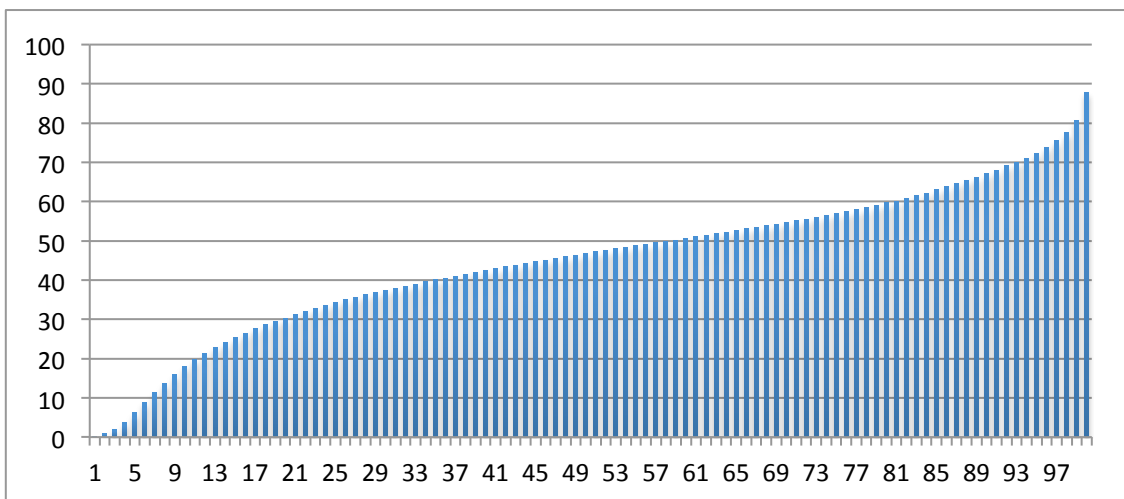
**Distribución de establecimientos según % de alumnos vulnerables. Umbral 11.734 puntos**



Fuente: Base de datos conjunta, FPS y RECH

**Gráfico 4:**

**Distribución de establecimientos según % de alumnos vulnerables. Umbral promedio ponderado**



Fuente: Base de datos conjunta, FPS y RECH

#### **4.- Distribución de establecimientos por dependencia y tamaño**

De especial interés resulta examinar la distribución de establecimientos según porcentaje de alumnos vulnerables a nivel de dependencia escolar y según tamaño de los establecimientos. El análisis se realiza para el umbral continuo de vulnerabilidad, por las razones citadas previamente.

El Cuadro 3 presenta la distribución por dependencia resumida a través de sus principales percentiles. Hay que tener en consideración que los percentiles descritos están referidos a las distribuciones de cada subgrupo. Así, el percentil 10 de los establecimientos municipales corresponde a un porcentaje de alumnos vulnerables del 35,4%; el percentil 10 de los privados subvencionados a un 16,3% y el percentil 10 de los particular pagados a un 0,2%. En todos los percentiles hay un mayor porcentaje de alumnos vulnerables en los establecimientos municipales que en los particulares subvencionados, y estos exhiben a su vez mayor vulnerabilidad que los particulares pagados.

En la última fila del Cuadro 3 se registra el promedio de alumnos vulnerables por tipo de dependencia: 51,5% en los municipales, 40,6% en los particular subvencionados y 3,5% en los particular pagados.

Es interesante hacer notar que la distribuciones de los establecimientos municipales y particular subvencionados difieren significativamente en la parte baja de la distribución (percentil 50 hacia abajo), mientras que tienden a parecerse en la mitad superior. Ello refleja la existencia de un conjunto de establecimientos particular subvencionados que posee un porcentaje relativamente bajo de alumnos vulnerables, mientras que para los establecimientos municipales la regla general es que alberguen un porcentaje más elevado de alumnos vulnerables. Sin embargo, el segmento de establecimientos particular subvencionados de elevada vulnerabilidad alberga porcentajes de alumnos vulnerables que son tan altos como los establecimientos municipales de mayor vulnerabilidad.

Por su parte, el Cuadro 4 presenta la distribución de establecimientos clasificados según su tamaño o número de alumnos matriculados. La información se presenta a nivel de quintiles de la matrícula y revela que hay una clara correlación inversa entre el nivel de vulnerabilidad y el tamaño del establecimiento: mientras más pequeños son los establecimientos, más probable es que sus alumnos presenten elevados grados de vulnerabilidad. Este resultado está fuertemente influenciado por la condición rural de los establecimientos más pequeños, pero sigue siendo vigente una vez que se controla por el área geográfica.

El 20% de establecimientos más pequeño es cualitativamente distinto. Este quintil está compuesto por aquellos establecimientos que tienen entre 1 y 20 alumnos, y su matrícula total representa tan solo a un 0,6% del universo de estudiantes. Sin embargo, tiene un peso significativo en la parte

superior de la distribución de establecimientos según niveles de vulnerabilidad. Ello puesto que cerca de tres quintas partes del 10% de los establecimientos con mayor porcentaje de alumnos vulnerables corresponden a escuelas pequeñas, con no más de 20 estudiantes.

**Cuadro 3: Distribución del porcentaje de alumnos vulnerables por dependencia educacional (umbral: promedio ponderado)**

	Dependencia			
	Municipal	Particular subvencionado	Particular pagado	Total
Percentil 10	35,4	16,3	0,2	19,1
Percentil 25	43,1	26,9	0,6	34,7
Percentil 50	51,1	39,9	1,9	47,0
Percentil 75	59,8	53,6	3,9	57,1
Percentil 90	69,0	65,8	8,0	67,5
Promedio	51,5	40,6	3,5	45,0

Fuente: Base de datos conjunta, FPS y RECH

**Cuadro 4: Distribución del porcentaje de alumnos vulnerables por quintiles de tamaño de establecimientos (umbral: promedio ponderado)**

	Tamaño establecimiento (quintiles de matrícula)				
	1-20	21-111	112-180	281-556	557-4478
Percentil 10	37,4	35,5	20,4	12,8	7,0
Percentil 25	48,3	46,3	36,2	28,6	21,9
Percentil 50	59,3	55,4	46,9	40,7	34,6
Percentil 75	69,9	64,1	54,1	49,3	43,6
Percentil 90	77,1	70,9	59,9	55,1	50,7
Promedio	58,2	53,7	43,4	37,6	32,1

Fuente: Base de datos conjunta, FPS y RECH

### 5.- Distribución de establecimientos bajo otros umbrales de vulnerabilidad

La elección del umbral continuo (promedio ponderado) para caracterizar socioeconómicamente a los establecimientos tiene fundamentos empíricos y conceptuales. No obstante, cabe preguntar cuán robusta es la clasificación socioeconómica de los establecimientos educacionales en comparación a la que entregan umbrales alternativos. Para examinar el punto se procede a

comparar el percentil que ocupa cada establecimiento en la distribución base versus el que obtiene en las distribuciones alternativas.

Los resultados se presentan en tramos según intervalos de percentiles. El primer tramo corresponde a desplazamientos hasta cinco percentiles en la dirección ascendente y hasta cinco percentiles en dirección descendente. Equivalentemente, el primer tramo representa desplazamientos dentro del decil de origen, si éste se define con centro en el establecimiento de referencia.<sup>7</sup> El segundo tramo corresponde a desplazamientos entre 6 y 10 percentiles, en dirección ascendente o descendente (se desplaza a un decil de distancia desde el origen). El tercer tramo se asocia a desplazamientos entre 11-15 percentiles (dos deciles de distancia) y el cuarto tramo agrupa a desplazamientos correspondientes a 16 o más percentiles (tres o más deciles distancia).

El Cuadro 5 presenta los resultados de la comparación de la distribución base con la asociada a 4.500 puntos como umbral de vulnerabilidad. Aquí se tiene que un 49% de los establecimientos experimenta desplazamientos dentro del decil de origen y otro 22,5% se desplaza a un decil de distancia. De manera que casi el 70% se mantiene dentro del quintil de origen (suma de los dos primeros tramos). Por otra parte, un 13% experimenta desplazamientos a dos deciles de distancia y un 15% se desplaza a tres o más deciles del origen.

El Cuadro 6 muestra que hay mayor estabilidad en la clasificación socioeconómica de establecimientos cuando se compara la distribución base con aquella que usa un umbral de 8.500 puntos. En este caso un 72% de los establecimientos permanece en el mismo decil de origen respecto de la distribución base, mientras que el porcentaje que se mantiene en el quintil de origen asciende a casi un 90%. Cuando la comparación se realiza respecto del umbral de 11.734 los respectivos porcentajes son 47,5% y 69%, un resultado que es más parecido al obtenido con el primer umbral (Cuadro 7).

Estos resultados muestran que la elección del umbral para definir vulnerabilidad tiene importancia, puesto que no da lo mismo operar con uno u otro umbral. Ello es especialmente cierto en el caso de los establecimientos municipales, que son los más proclives a modificar su posición en el ordenamiento socioeconómico cuando cambia el umbral. Parte de esta mayor movilidad se asocia a las escuelas más pequeñas, donde cualquier medida de distribución tiende a ser más inestable. También importa el nivel de vulnerabilidad; cuando este es bajo el establecimiento tiende a permanecer en su posición original, porque los cambios en los umbrales no afectan el indicador (como sucede con el segmento particular pagado).<sup>8</sup>

Más allá de las explicaciones puntuales existe una interpretación general detrás de los resultados descritos. El análisis que se realiza se refiere la relación entre el ranking de establecimientos y el

---

<sup>7</sup> Por ejemplo, un establecimiento en el percentil 15 en la distribución base se mantiene en el decil de origen si se sitúa entre los percentiles 10 y 20 de la distribución alternativa.

<sup>8</sup> Si todos los alumnos de un establecimiento tienen más 13.000 puntos en la FPS, el porcentaje de vulnerabilidad será 0% para cualquiera de los umbrales analizados.

parámetro (umbral) que se usa para caracterizar la vulnerabilidad al interior de cada establecimiento. Sin embargo, la vulnerabilidad depende de la distribución de la variable socioeconómica entre los alumnos de cada establecimiento. Que el ranking de establecimientos sea sensitivo al umbral para definir vulnerabilidad refleja la dificultad inherente a representar en un solo parametro (umbral) la información contenida en la distribución de alumnos al interior del establecimiento. El asunto de fondo es que un solo parametro no puede captar toda la información que hay en la distribución, a menos que esta tenga una forma funcional muy sencilla.

Este es un resultado general, que trasciende a los indicadores analizados en el documento. A modo de ejemplo se puede citar el caso de la medición de la desigualdad de ingresos. Hay un conjunto de parametros que miden la desigualdad: coeficiente de Gini, índice de Theil, razón de quintiles, etc. Ninguno de ellos por sí solo es capaz de captar toda la información contenida en la distribución de ingresos. Por ejemplo, el coeficiente de Gini resulta ser especialmente sensible a la información contenida en la parte media de la distribución mientras que el índice de Theil captura mejor la información de los extremos. De esta manera, los rankings de desigualdad entre los países son sensitivos al parámetro que se está ocupando para ordenar a los países de mayor a menor desigualdad. Chile puede estar mejor o peor clasificado dentro de América Latina dependiendo del indicador usado para medir la desigualdad de ingresos.

De la discusión anterior sería erróneo concluir que es imposible realizar rankings, sino que el indicador utilizado importa y por tanto debe elegirse de acuerdo a sus atributos intrínsecos.

Cuadro 4: Desplazamientos entre distribución base y umbral de 4.500 puntos (% establecimientos)

Percentiles de desplazamiento (ascendente o descendente)	Municipal	Privado subv	Particular pagado	Total
0-5	43,0	52,0	97,3	49,0
6-10	23,1	24,1	1,1	22,5
11-15	14,0	13,5	0,7	13,2
16 y más	19,9	10,4	0,9	15,4
Total	100,0	100,0	100,0	100,0

Fuente: Base de datos conjunta, FPS y RECH



Cuadro 5: Desplazamientos entre distribución base y umbral de 8.500 puntos (% establecimientos)

Percentiles de desplazamiento (ascendente o descendente)	Municipal	Privado subv	Particular pagado	Total
0-5	63,6	81,5	99,6	72,1
6-10	21,8	12,9	0,2	17,4
11-15	7,3	3,8	0,0	5,6
16 y más	7,3	1,9	0,2	4,9
Total	100,0	100,0	100,0	100,0

Fuente: Base de datos conjunta, FPS y RECH

Cuadro 6: Desplazamientos entre distribución base y umbral de 11.734 puntos (% establecimientos)

Percentiles de desplazamiento (ascendente o descendente)	Municipal	Privado subv	Particular pagado	Total
0-5	36,0	58,1	99,3	47,4
6-10	25,0	19,2	0,2	21,6
11-15	17,1	10,7	0,5	13,9
16 y más	21,9	12,1	0,0	17,1
Total	100,0	100,0	100,0	100,0

Fuente: Base de datos conjunta, FPS y RECH

## 6.- Cambios a futuro en la Ficha de Protección Social

En su actual versión la Ficha de Protección Social asigna un puntaje al hogar en base a la capacidad de generación de ingresos que tienen los miembros adultos del hogar, que a su vez es una función de variables como los años de escolaridad, la experiencia laboral, la situación ocupacional y las características del mercado laboral en la zona de residencia, entre otras. El modelo asume que no poseen capacidad de generación de ingresos las mujeres embarazadas, las madres de menores de dos años, los discapacitados, los estudiantes y los adultos mayores. El puntaje de la Ficha también considera ingresos del hogar que no provienen del trabajo, como subsidios y pensiones. El conjunto de los recursos económicos es dividido por un índice de necesidades del hogar, que

considera el número y composición de los miembros así como la existencia de personas con necesidades adicionales (discapacitados).

La Ficha de Protección Social entró en funcionamiento en el año 2006, cuando reemplazó a la Ficha Cas que hasta entonces servía para seleccionar a los usuarios de los programas sociales focalizados. Actualmente, se encuentra bajo un proceso de revisión puesto que se ha detectado un significativo grado subdeclaración en los datos socioeconómicos reportados por una parte de los hogares entrevistados.

A futuro se prevé que todas las variables que considere la nueva versión de la Ficha provengan de registros administrativos o puedan ser objeto de verificación por parte de los organismos responsables. Al mismo tiempo, Mideplan está fortaleciendo de modo sustancial su capacidad de monitoreo y fiscalización del instrumento. El objetivo es que la Ficha en su versión revisada no pueda ser manipulada por parte de los encuestados, de manera que entregue información confiable y fidedigna sobre la condición socioeconómica de los hogares.

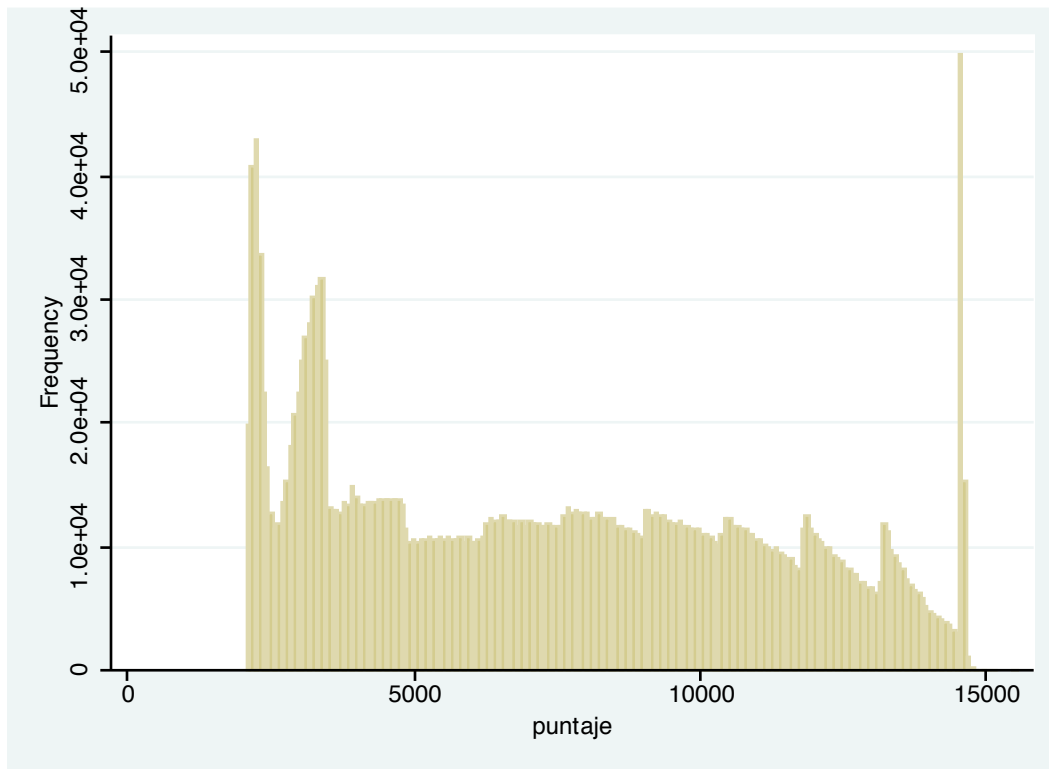
Se prevé que la nueva Ficha entre en operación hacia noviembre del 2011 y que la renovación de la información para el conjunto de la población encuestada tarde aproximadamente un año. De modo que hacia fines del 2012 se contaría con puntajes revisados del instrumento. La corrección debiera causar un incremento en el puntaje promedio, puesto que ajusta hacia arriba las estimaciones de la condición socioeconómica que han sido subdeclaradas en el actual instrumento. También se espera una reordenación de los hogares en la distribución de puntajes, desplazándose en la dirección ascendente los casos de subdeclaración.

La pregunta relevante a efectos del presente análisis es si habrá cambios de importancia en la clasificación socioeconómica de los establecimientos. En principio, una elevación de puntajes de la Ficha que se distribuya en forma relativamente pareja en la población no debiera afectar la ordenación o ranking de los establecimientos; puesto que todos ellos tenderían a reducir de modo aproximado el porcentaje de alumnos vulnerables.

Sin embargo, pudiera haber modificaciones sistemáticas en la clasificación de establecimientos porque la recolección de datos para la actual la ficha fue realizada con criterios disímiles por parte de algunos de los equipos municipales. Ello provocaría cambios en los puntajes con efectos asimétricos según comunas, en cuyo caso cabría esperar re-ordenaciones en la distribución que se basa en el porcentaje de alumnos vulnerables. El Anexo C contiene un análisis en la materia.

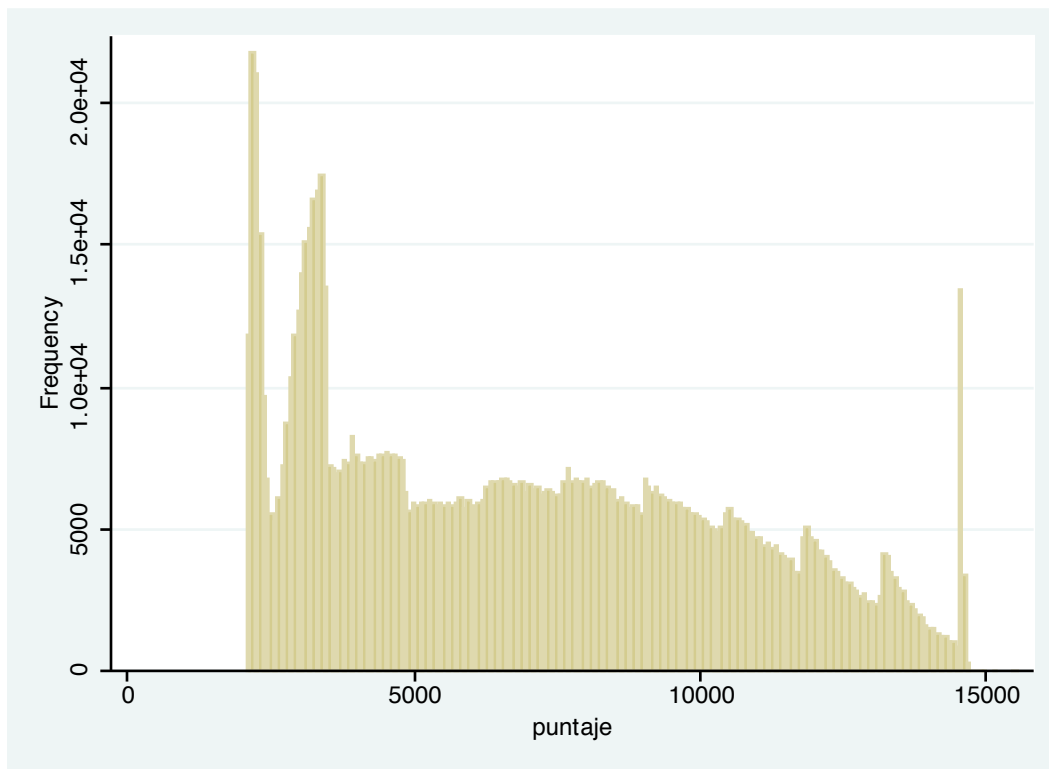
## Anexo A

Gráfico A-1: Distribución de puntajes de la Ficha de Protección Social en la población de estudiantes de enseñanza básica y media, 2010



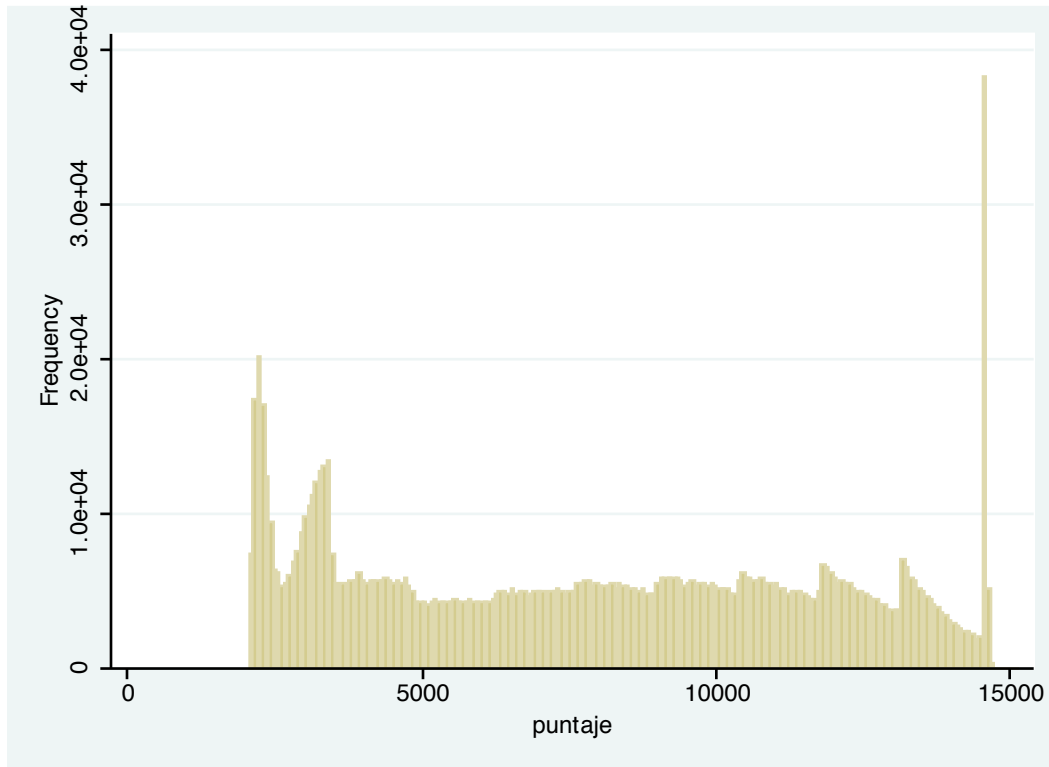
Fuente: Base de datos conjunta FPS y RECH

Gráfico A-2: Distribución de puntajes de la Ficha de Protección Social en la población de estudiantes de enseñanza básica y media, establecimientos municipales, 2010



Fuente: Base de datos conjunta FPS y RECH

Gráfico A-3: Distribución de puntajes de la Ficha de Protección Social en la población de estudiantes de enseñanza básica y media, establecimientos particular subvencionados, 2010



Fuente: Base de datos conjunta FPS y RECH

## **Anexo B: Estabilidad de la clasificación Simce**

La clasificación socioeconómica de establecimientos basada en la información del Simce no es estable en el tiempo, sino que está afecta a cambios tal como se muestra en el Cuadro B-1 para el período 2005 a 2009. Allí se muestra el porcentaje de establecimientos que cambia de grupo socioeconómico entre el año 2005 y cada uno de los años siguientes, así como aquel que cambia de grupo al menos una vez en el período. El número de establecimientos que cambia de grupo de un año a otro es significativo, entre un quinto y un tercio del total, siendo mayor el porcentaje de cambio a medida que aumenta el intervalo de comparación. Por su parte, el porcentaje de establecimientos que cambia de grupo al menos una vez en todo el período (última fila) es mayor que cualquiera de las comparaciones bianuales, un resultado esperado, aún cuando este porcentaje parece ser menor a la sugeriría la acumulación de cambios en el tiempo.

Los cambios de clasificación socioeconómica por parte de los establecimientos pueden responder a dos tipos de factores: (i) estructurales, vinculados a modificaciones sistemáticas en la composición social de los estudiantes; (ii) aleatorias, asociadas a factores que empujan en una u otra dirección, sin que existe una razón sistemática subyacente. La aleatoriedad es una característica que se quisiera minimizar puesto que es un “ruido estadístico” que distorsiona la medida y origina un exceso de movilidad en el tiempo. La aleatoriedad se origina en errores no sistemáticos en la recolección de los datos (cuando los padres llenan los cuestionarios del Simce) y en el carácter muestral de los datos, puesto que los alumnos de cuarto básico de un determinado año calendario representan un estimador impreciso de la composición de la población total del establecimiento.

Hay que notar que la aleatoriedad causa inestabilidad tanto en la clasificación de establecimientos al interior de los grupos Simce, como en la propia construcción de los grupos a través de la técnica de conglomerados. Adicionalmente, la inestabilidad de la clasificación se amplifica con la utilización de grupos con límites fijos, donde la probabilidad de cambio de grupo depende de la distancia que exista entre los establecimientos y el umbral de cada grupo.

Una aproximación a la “movilidad estructural” viene dada por el incremento en el porcentaje de establecimientos que cambian de clasificación a medida que aumenta el período de comparación, asumiendo que el porcentaje de cambio aleatorio debiera mantenerse relativamente constante entre pares de años. De este modo, se podría inferir que alrededor de una séptima parte del cambio registrado en el cuadro B-1 se debe a movilidad estructural y seis séptimos a factores aleatorios (el incremento de cambio entre 2005/06 y 2005/09 entrega un aumento anual promedio cercano a 3 puntos porcentuales, que representan la séptima parte del cambio entre 2005 y 2006).

Debe también notarse que el porcentaje de establecimientos que cambia de categoría al menos una vez en todo el período es mayor a cualquier comparación entre pares de años, un resultado esperado, pero bastante menor a la suma de las partes. Esto último se debería a que la mayor parte de los cambios ocurre en un número acotado de establecimientos, que son aquellos que están más cerca de los umbrales en los grupos socioeconómicos de pertenencia.

**Cuadro B-1: Porcentaje de establecimientos que cambian de grupo Simce (4 básico), 2005 a 2009**

	A	B	C	D	E	Total
2005/06	18,0	24,0	19,4	19,7	9,8	19,9
2005/07	13,0	26,4	25,3	26,3	9,8	21,7
2005/08	21,6	29,5	28,5	22,5	8,3	24,9
2005/09	30,4	32,7	35,1	36,6	20,7	32,5
cualquier año	46,4	57,4	46,4	46,2	23,2	48,3

Fuente: Base de datos pruebas Simce, años respectivos

### **Anexo C. Posibles cambios en la clasificación de comunas en nueva Ficha de Protección Social**

A fines ilustrativos se realiza el ejercicio de comparar el puntaje promedio efectivo de la FPS en cada comuna versus el puntaje promedio esperado según la información de la encuesta Casen. El primer dato proviene de los registros administrativos de la FPS y ha sido provisto por Mideplan. El segundo dato proviene de una simulación del puntaje que los hogares obtendría en la Ficha en base a la información declarada en la encuesta Casen, que en principio no está afectada a los problemas de subdelcaracion presentes en la FPS. La simulación de la FPS en la Casen fue también realizada por Mideplan.

El ejercicio requiere dos ajustes previos. Primero, el promedio de la FPS en cada comuna que se simula en la Casen se realiza sobre la base de la cobertura observada en la FPS, suponiendo que son los hogares de menor nivel de ingreso per capita quienes acceden a la Ficha. Por ejemplo, si en la comuna de Coquimbo hay un 70% de hogares con Ficha, la simulación en la Casen se realiza considerando al 70% de hogares de menor ingreso per capita. Segundo, para reducir el error asociado a tamaños de muestras pequeñas se trabaja en forma conjunta con las encuestas Casen 2006 y 2009, de modo que se cuenta con aproximadamente el doble de observaciones muestrales que en caso de trabajar con la Casen 2009.

Los resultados del ejercicio se presentan en el Cuadro C-1, que identifica a las comunas que presentan la mayor brecha (sobre 25%) entre el puntaje promedio efectivo de la FPS y el puntaje simulado en la Casen. En principio, estas serían las comunas que debieran presentar las mayores modificaciones cuando se tenga los puntajes revisados de la FPS, esperándose aumentos en su nivel socioeconómico por arriba del promedio. Se insiste que el ejercicio tiene fines ilustrativos.

Cuadro C-1: Comunas que presentan mayores diferencias entre puntaj FPS efectivo vs simulado

<b>Código</b>	<b>Nombre</b>	<b>Gap</b>
13303	Tiltil	-0.3980912
13126	Quinta Normal	-0.3889028
6108	Machalí	-0.3644404
13301	Colina	-0.3630598
13129	San Joaquín	-0.3619758
13404	Paine	-0.3574553
8314	Alto Bío Bío	-0.3360367



7109	San Clemente	-0.3287807
5504	La Cruz	-0.3208422
8414	Quirihue	-0.3142626
7107	Pencahue	-0.3137577
5303	Rinconada	-0.3063945
9121	Colchol	-0.3061721
10108	Mauñín	-0.30485
10202	Ancud	-0.3013607
6112	Peumo	-0.2935646
3102	Caldera	-0.2898536
13130	San Miguel	-0.2867975
13401	San Bernardo	-0.2848795
10207	Queilén	-0.2812933
13201	Puente Alto	-0.2801085
14107	Paillico	-0.2769839
13106	Estación Central	-0.2766324
13109	La Cisterna	-0.2752074
10307	San Pablo	-0.2735744
10305	Rio Negro	-0.2729113
3301	Vallenar	-0.2726875
5302	Calle Larga	-0.2703662
10107	Llanquihue	-0.2694321
13124	Pudahuel	-0.2689893
8416	San Carlos	-0.268909
10301	Osorno	-0.2674704
13121	Pedro Aguirre Cerda	-0.2651166
6102	Codegua	-0.2599294
7404	Parral	-0.2589135
6103	Coinco	-0.2575482
15101	Arica	-0.2566165
8419	San Nicolás	-0.2561223
13402	Buín	-0.2548471
8406	Chillán Viejo	-0.2545122
13503	Curacaví	-0.2530885
13122	Peñalolén	-0.252832
10101	Puerto Montt	-0.2526312
5703	Lay Llay	-0.252475

Fuente: Encuesta Casen 2006 y base de datos de la FPS.

