



DOCUMENTOS DE TRABAJO

Del PROYECTO PICT 2003 N° 14534

FONCYT / SECYT

DOCUMENTO DE TRABAJO N° 2 / 2005

Un Índice de nivel socio-económico de los hogares para aplicar a un modelo explicativo del rendimiento escolar

Verónica Herrero, Paula Palacios y María Florencia Ruiz Díaz

Octubre 2005

INSTITUTO DE ECONOMÍA Y FINANZAS

FACULTAD DE CIENCIAS ECONÓMICAS

UNIVERSIDAD NACIONAL DE CÓRDOBA

**UN INDICE DE NIVEL SOCIO-ECONÓMICO DE LOS HOGARES PARA APLICAR A UN MODELO
EXPLICATIVO DEL RENDIMIENTO ESCOLAR**

Verónica Herrero, Ana Paula Palacios, María Florencia Ruiz Díaz

Septiembre 2005

Facultad de Ciencias Económicas
Universidad Nacional de Córdoba

Para ser presentado al XXXIII Coloquio Argentino de Estadística. Villa Giardino, Córdoba, 5 al 7 de
Octubre de 2005

UN INDICE DE NIVEL SOCIO-ECONÓMICO DE LOS HOGARES PARA APLICAR A UN MODELO EXPLICATIVO DEL RENDIMIENTO ESCOLAR

Verónica Herrero, Ana Paula Palacios, María Florencia Ruiz Díaz
Facultad de Ciencias Económicas. Universidad Nacional de Córdoba

RESUMEN

El nivel socio-económico de los hogares de los estudiantes es uno de los factores que aparece en investigaciones del campo de la educación como determinante del rendimiento escolar del alumno. En el marco de un estudio sobre la calidad de la educación en Argentina, los autores del presente estudio tuvieron la necesidad de contar con un índice que capte esa variable para los alumnos de la población bajo análisis, con la finalidad de incorporarla a un modelo explicativo que incluye un conjunto más amplio de factores. Dicho índice fue construido empleando las variables disponibles de capital del hogar en la base de datos correspondiente al Operativo Nacional de Evaluación de la Calidad Educativa 2000, entre las que no figura el ingreso de la familia que suele ser utilizado como una primer variable proxy de nivel socio-económico ni la ocupación del principal sostén del hogar. Este trabajo reporta la metodología utilizada, las limitaciones existentes y presenta los índices obtenidos para el país y sus regiones. Componentes principales y análisis factorial de correspondencias fueron herramientas que ayudaron en la confección del índice. Finalmente se cotejan los resultados con los obtenidos al replicar el cálculo sobre una población diferente, correspondiente a una muestra oficial nacional que contiene además información sobre ingreso del hogar.

PALABRAS CLAVES: índices socio-económicos, hogar, estudiante, nivel de vida, capital del hogar, rendimiento escolar.

UN ÍNDICE DE NIVEL SOCIO-ECONÓMICO DE LOS HOGARES PARA APLICAR A UN MODELO EXPLICATIVO DEL RENDIMIENTO ESCOLAR

Verónica Herrero, Ana Paula Palacios, María Florencia Ruiz Díaz⁺
Facultad de Ciencias Económicas. Universidad Nacional de Córdoba

1. La dimensión socio-económica como determinante del rendimiento escolar ¿Por qué estudiar la relación?

Los modelos explicativos del rendimiento escolar generalmente incorporan variables escolares, entre las que se cuentan características del establecimiento educativo, capacidad del docente, actitud del grupo de alumnos que integran el curso y variables que reflejan el efecto de factores externos a la escuela, tales como nivel educativo de los padres, actitud de los padres hacia el estudio de los hijos, calidad del hogar y otros que integran el concepto de nivel socio-económico del hogar (Levin, 1971, Monk, 1989, Rowe, 1995, Braun y Mirlevy, 2005). El propósito de este trabajo es presentar resultados preliminares de un índice de nivel socio-económico del hogar al que pertenecen los alumnos, que fue confeccionado con el fin de ser incorporado a un estudio de rendimiento escolar en la Argentina.

Asimismo, es importante, en el marco de modelos explicativos del rendimiento escolar, conocer el comportamiento de la variable nivel socio- económico de los hogares de los alumnos para tener la posibilidad de:

- a) Cuantificar el efecto de los otros factores que inciden en el rendimiento escolar, cuando se controla la variable nivel socio- económico.
- b) Detectar diferencias de rendimientos según el contexto en el que el nivel socio- económico esté enmarcado (ámbito regional), teniendo en cuenta que existe un contexto general (país). En este sentido, determinadas carencias en el hogar de un estudiante tienen diferente incidencia en el rendimiento, según la zona del país en que vive (vbgr. el efecto de la falta de calefacción en una vivienda del sur del país es diferente respecto a otra ubicada en algunas provincias del norte).

Estas consideraciones reflejan la importancia de contar con un índice de nivel socio-económico (INSE) que pueda actuar como variable independiente dentro de un modelo explicativo del rendimiento escolar individual de los estudiantes. La construcción de un índice de nivel socio-económico presenta complejidades importantes asociadas con la selección del conjunto de factores y el peso asignado a cada uno. Típicamente, en este tipo de trabajos se consideran tanto el nivel de educación como la condición de ocupación de los padres, la posesión de distintos bienes y las principales características de la vivienda.

Este trabajo propone un índice de nivel socio-económico que se apoya en tres indicadores básicos: nivel de estudios del padre, condición de hacinamiento de la vivienda y posesión de ciertos bienes. La validez del indicador puede verse en términos de la adaptación específica del constructo al concepto teórico que quiere reflejar (validez sustantiva), como también en cuanto a la consistencia interna que evidencia la capacidad para ordenar adecuadamente los hogares de los alumnos con diferente grado de manifestación de los ítems considerados –validez estructural- (en este caso de un índice cuantitativo). Además se espera que el índice sea aplicable a otros datos análogos y mantenga su capacidad de jerarquización de situaciones –propiedad conocida como validez externa- (Clark y Watson, 1995).

El índice es calculado a partir de la información proporcionada por los Operativos Nacionales de Evaluación de la Calidad Educativa (ONE) que realiza el Ministerio de Educación, Ciencia y Tecnología de la Nación desde el año 1993. Además de las pruebas de rendimiento escolar receptadas con motivo de esos operativos, cada alumno debe responder a un conjunto de preguntas, dentro de un cuestionario auto-administrado, las que sirvieron de base para confeccionar el índice de nivel socio-económico que se reporta en este trabajo.

En la siguiente sección se examinan los antecedentes existentes respecto del tema. Luego, se efectúa una descripción de la metodología aplicada para el cálculo del INSE indicando además las características de los datos disponibles y las variables tomadas en cuenta para su cómputo.

⁺El trabajo forma parte del Proyecto SECYT-PICT2003 N° 14534. El trabajo fue comentado y discutido por el equipo del proyecto en distintos seminarios internos del IEF. Se agradecen los comentarios de: Dr. Héctor Gertel, Dr. Roberto Giuliadori, Dra. Mariana De Santis, Lic. Daniela Cristina, Lic. María Luz Vera, Diego Fresoli y Guadalupe Morra, integrantes del equipo.

Posteriormente se presentan los resultados de dos índices computados, a partir de la consideración simultánea de las variables de interés (haciendo uso de elementos surgidos tras aplicar la técnica de Análisis Factorial de Correspondencias Múltiples y Componentes Principales). Finalmente, con el propósito de contar con un elemento externo de validación de la medida propuesta, se aplica uno de los índices computados, a una muestra de hogares más general, representativa del total nacional, y se comparan los resultados (Encuesta de Condiciones de Vida, Siempre 2001). La sección final contiene una síntesis de los principales resultados.

2. Antecedentes

La literatura sobre estratificación socio-económica elaborada a partir de estadísticas sociales, en el país, gira de manera predominante en torno a la clasificación de los hogares entre pobres (indigentes y el resto de la población pobre, por ejemplo) y no pobres¹. Son menos frecuentes los estudios, que a partir de información de las distintas manifestaciones de capital de los hogares, han tratado de encontrar grados diferenciados de status socio-económico.

Para estudiar la relación entre rendimiento escolar y nivel socio-económico de los hogares tal clasificación dicotómica resulta insuficiente, de ahí la necesidad de contar con un índice que permita aproximar la diversidad de situaciones de los hogares. Un primer esfuerzo por cubrir esta necesidad se presenta en Llach y Montoya (1999) quienes calcularon índices sencillos a partir de las Evaluaciones de Calidad Educativa del año 1993. Dicho índice es un promedio simple de la posesión de diferentes bienes en el hogar y el máximo nivel de escolaridad alcanzado por los padres. Con relación a los diferentes niveles de escolaridad se imputan valores numéricos que varían de 1 a 6 (desde 1 en el caso de no poseer estudios hasta 6 para superior completo o incompleto). Con respecto al Indicador de Posesión de Bienes, construyen un índice basado en el número de bienes que los niños poseen en su hogar (entre los 16 incluidos en el cuestionario). Este indicador, por lo tanto varía entre 1 y 16.

Otro antecedente es el informe vinculado con las brechas socioeducativas entre las provincias (CIPPEC, 2004), donde se trabaja a partir de la información de los hogares de los alumnos recogida en oportunidad de las pruebas del ONE. Allí se presenta un indicador de nivel socio-económico con cuatro componentes: nivel educativo del padre y de la madre, hacinamiento, condiciones sanitarias de la vivienda y posesión de bienes, éstos últimos con ponderaciones diferenciadas entre sí de acuerdo con una clasificación entre bienes básicos, no básicos y de lujo.

Dentro de América Latina, en un trabajo con datos de las pruebas educativas desarrolladas en Perú (Caro, 2002), incorpora a la formulación de un índice estándar, el cómputo del primer factor (usando Componentes Principales), con un conjunto amplio de variables relativas a las condiciones de la vivienda, la posesión de bienes y otras disponibles, que sintetizarían el nivel socio-económico de los alumnos. Además, añade como método de control de los resultados obtenidos, el contraste de éstos con variables como el gasto de los hogares, a fin de asegurarse la consistencia externa del índice creado.

Este tipo de índice de nivel socio-económico ha sido encontrado también en estudios vinculados con el efecto de la calidad del entorno familiar sobre el desarrollo del niño, y en particular en el rendimiento escolar (como en Veguer y Serra, 1996), concentrado en la situación de los menores en España. Allí se han utilizado como indicadores de situación socio-económica del hogar la escolaridad y ocupación del jefe de familia. En particular, se le asigna una influencia negativa en el índice a la situación de desocupación del cabeza de familia.

Con otros propósitos la literatura presenta esfuerzos por vincular los índices de status con escalas de ingresos. La Asociación Argentina de Marketing construyó en el año 2002 un INSE que fue actualizado en 2004. La primera versión era un índice de nivel socio-económico de uso generalizado que buscaba clasificar hogares según sus características relevadas a partir de un conjunto de encuestas. La información considerada incluía el máximo nivel educativo alcanzado por el principal sostén del hogar (persona que más dinero aporta), el acceso a cobertura de salud, disponibilidad de tarjeta de débito, computadora e Internet en el hogar; categoría ocupacional del principal sostén del hogar y número de aportantes². En la revisión realizada en 2004, se aplicaron a la misma base de

¹ Por ejemplo, Coudouel y otros, 2002; Gómez y otros, 2003; INDEC, 2003; Lee, 2003; Minujín, 1992; Perry, 2002.

² Las ponderaciones resultantes fueron: categoría ocupacional del PSH (32); cantidad de autos (22); posesión de PC, Internet y tarjeta de débito (19); educación del PSH (13); número de aportantes al hogar (9); cobertura médica (5).

datos del estudio de 2002, técnicas de Análisis Factorial de Correspondencias Múltiples (AFCM), Clusters y Análisis Discriminante, con la finalidad de mejorar la caracterización de la estratificación social del país obteniéndose 6 grupos socio-económicos (homogéneos internamente, heterogéneos entre sí), a los que denominaron: Alta; Media Competitiva; Media Empobrecida; Baja; Marginal y Pasiva³.

Para Nueva Zelanda, Davis, Jenkin y Coope (2002) construyeron índices de status socio-económico a partir de datos censales, donde la variable clave considerada es la ocupación. La información disponible en ese caso les permitió asignar puntajes específicos a cada ocupación según las premisas del modelo de retornos al capital humano, en base a la cual se construyó un índice de estratificación social. Así, la estructura ocupacional es el nexo que vincula la relación fundamental entre capital cultural (educación) con el acceso a retornos materiales (ingreso).

En suma la construcción de índices de nivel socioeconómico persigue propósitos múltiples: en algunos casos la preocupación es identificar hogares pobres, en otras construir una escala que refleje la diversidad de situaciones de los hogares y finalmente, estudiar la vinculación entre la situación socio-económica y los ingresos de los hogares. En el caso específico de los Operativos Nacionales de Evaluación no existe información sobre ingresos por lo que este trabajo, al igual que al de Llach y Montoya citado anteriormente, presenta una caracterización de los hogares basada solamente en variables que reflejan el stock de capital físico y de capital humano. Pero a diferencia del índice simple que se expone en dicho trabajo, aquí se construyen y discuten distintas alternativas.

3. Fuentes de datos

El Operativo Nacional de Evaluación de la Calidad Educativa correspondiente al año 2000 es la fuente de datos utilizada. Si bien el Operativo viene desarrollándose desde 1993 con carácter muestral, solo en el año 2000 se relevó información para el total de la población objetivo. Para este estudio se seleccionaron los alumnos correspondientes al 6to grado o del EGB. Al momento del diseño de este trabajo, los datos correspondientes al año 2000 eran los últimos disponibles. No obstante, hay algunos elementos que constituyen una excepción a la cobertura universal: en primer lugar, el Operativo no incluyó a los establecimientos de la provincia de Neuquén y en segundo término, tampoco forman parte del relevamiento los establecimientos rurales con secciones de menos de 5 alumnos.

Las bases de datos del operativo han sido presentadas agrupando la información en tres grupos a saber: alumnos, docentes y directivos. Estos grupos se corresponden con bases de datos independientes existiendo un código que permite vincularlas.

Para este trabajo se utilizará solamente la base de alumnos la cual contiene información correspondiente al resultado de la evaluación de cada alumno en las asignaturas de Matemática y Lengua, así como datos provenientes de una encuesta que se administra de manera adicional (Cuadernillo Complementario del Alumno). El Cuadernillo contiene información acerca de las características sociodemográficas del alumno, los miembros de su hogar y la vivienda donde reside, la trayectoria escolar, opiniones sobre el ambiente familiar y escolar. A partir de este cuerpo de variables se trabajó para obtener los índices de nivel socio-económico objeto del presente estudio.

La base de datos de alumnos del 2000 contiene 613,919 casos, con la distribución por regiones que se detalla en la Tabla I y que corresponde a las regiones definidas por el INDEC. Una vez depurada la base, resultan 263,767 casos (43% de la base total) con información completa para el conjunto de variables de interés.

³ Las variables seleccionadas entre las disponibles y los pesos resultantes del análisis de regresión para predecir ingresos fueron: Miembros del Hogar – Uno (19); Ocupación del PSH- Dirección (15); TV Satelital – Si (13); Miembros del Hogar – Dos (11); Ocupación del PSH - Con empleados (11); Tiene empleada doméstica – Si (9); Tres o más Baños (7); DVD reproductor – Si (6); Educación del PSH – Universitario completo (5); Educación del PSH – Secundario completo (4); TV Cable – Si (4); Aire

Tabla I: Distribución de los alumnos evaluados por región

Región	Población relevada de alumnos en 6º grado		Muestra de trabajo (con información completa sobre las variables de análisis)	
	Total	Porcentaje	Total	Porcentaje
Capital Federal	36,713	5.98	17,473	6.62
GBA	154,543	25.17	58,621	22.22
Pampeana	208,906	34.03	92,275	34.98
NOEA	78,972	12.86	33,957	12.87
NEA	63,072	10.27	26,763	10.15
Cuyo	45,400	7.4	22,502	8.53
Patagonia	26,313	4.29	12,176	4.62
Total	613,919	100	263,767	100.00

Fuente: Elaboración propia con datos de ONE 2000 (Base de Alumnos)

La tabla I permite verificar que la distribución regional de alumnos con información completa (263,767) resulta similar a la distribución del total de alumnos participantes en el Operativo (613,919).

4. Metodología empleada

a. Variables consideradas y operacionalización

- *Nivel de instrucción del padre*

En la literatura se considera una variable proxy de la capacidad de obtener ingresos del hogar. En otros estudios se ha considerado también la educación de la madre (Llach, 1999, CIPPEC, 2004). Para el cálculo del INSE sólo se tomará en cuenta la instrucción del padre como indicador del capital humano del hogar. En particular en los estudios de referencia acerca de los determinantes del rendimiento escolar, la educación de la madre se suele privilegiar como un control acerca de la capacidad para motivar al alumno y para asistirlo en la realización de las tareas vinculadas con el aprendizaje, más que como indicador del nivel socio-económico de los hogares. Sin embargo, se destacan algunas limitaciones vinculadas con la variable nivel de instrucción del padre:

- No disponemos de información acerca de si el padre reside o no en el mismo hogar que el alumno.
- En algunos hogares el principal aportante de ingresos puede ser la madre u otro miembro del hogar.

- *Posesión de bienes en el hogar*

El interés sobre este tipo de variables radica en que la posesión de mayor cantidad de bienes es una exhibición de mayor riqueza. Las principales limitaciones de las variables disponibles se vinculan con:

- No se dispone de información acerca de la calidad y/o antigüedad del bien, lo cual reflejaría mejor el valor económico del mismo.
- Algunos bienes son relativamente básicos (satisfacen necesidades elementales) y otros, necesidades probablemente secundarias –si bien esta valoración también es relativa a cada individuo-.
- La posesión de determinados bienes está altamente difundida en la población por lo cual la posesión del mismo no parece un buen discriminador de NSE.

- *Condición de hacinamiento*

Para calcular la condición de hacinamiento se construyó el índice de hacinamiento, el cual se define como el cociente entre la variable Cantidad de miembros y Cantidad de ambientes, y el indicador de referencia se obtiene como la recíproca de este índice (debido a que, en este caso un valor pequeño del cociente indica una mejor situación relativa). Entre las limitaciones vinculadas con las variables disponibles pueden mencionarse:

- Las variables agrupan en la última categoría diferentes cantidades de personas y ambientes. Se utilizó el valor del límite inferior del intervalo.

acondicionado – Si (4); Medicina Prepaga del PSH o cónyuge (4); Ocupación del PSH - Sin empleados (4); Miembros del Hogar - Tres a Cuatro (3); Tarjeta de crédito – Si (3).

- La variable número de ambientes (no se indagó acerca de habitaciones o cuartos de uso exclusivo) no corresponde a dormitorios, por lo tanto no medimos exactamente el mismo concepto que con los indicadores habituales de hacinamiento.

- *Agua corriente, Instalación eléctrica y Uso de habitación para cocina y cuarto*

Estas variables no fueron incluidas debido a que posiblemente sólo aportaban información para diferenciar la situación de los individuos en situación de extrema pobreza del resto, y no para posibilitar la clasificación de individuos en diferentes grados.

b. Método utilizado

Se han elaborado tres diferentes versiones del *INSE*: $INSE_1$, $INSE_2$ e $INSE_3$. En los tres casos el índice obtenido posee los siguientes componentes, asociados con distintas manifestaciones de capital del hogar:

- Educación del padre
- Posesión de bienes durables en el hogar
- Indicador de hacinamiento

Los tres componentes corresponden a índices independientes, con un rango de valores entre 0 y 100 cada uno, que son promediados para obtener el *INSE* global.

$$INSE_k = \sum_{i=1}^3 w_i I_i \quad (1)$$

donde w_i es la ponderación e I_i es el índice correspondiente a cada componente de la manifestación del capital del hogar: $I_1=IE$ (Índice de educación del padre); $I_2=IH$ (Índice de hacinamiento); $I_3=IBD$ (Índice de bienes durables).

Cada índice generado a nivel de cada alumno, se obtuvo a partir de la siguiente expresión:

$$I_{i_n} = \frac{[x_{i_n} - \text{Min}_j(x_i)]}{[\text{Max}_j(x_i) - \text{Min}_j(x_i)]} \cdot 100 \quad (2)$$

donde para cada individuo n , $i=1, 2, 3$; $j=1, \dots, 263767$; x representa la cuantificación de la expresión de capital del hogar correspondiente.

En el caso de la educación del padre, los valores posibles de x varían entre 1 y 6, donde 1 corresponde al mínimo de escolaridad (sin instrucción o con primario incompleto) y 6 al máximo detallado en la encuesta (superior –terciaria o universitaria- completa).

Para el indicador de hacinamiento, los valores posibles de x surgen de la siguiente función calculada para cada individuo:

$$x_{2n} = 1/(q_n / h_n) \quad (3)$$

Donde para el individuo n , q corresponde al total de personas que residen en el hogar del alumno, y h el total de habitaciones. La expresión contenida en el denominador es el indicador conocido como índice de hacinamiento.

Los valores mínimos y máximos respectivamente de x para los diferentes individuos de la base, resultaron: 0.083 (correspondientes a una relación de personas por habitación de 12) y 3.5 (derivado de una relación de personas por habitación de 0.29), respectivamente.

En el caso de la posesión de bienes durables, en primer lugar se procedió a la suma de bienes del hogar de cada alumno. Pueden adoptarse distintos criterios para la ponderación de los bienes en la suma. En general la fórmula (4) especifica esta construcción.

$$IBD_{kn} = I_{3kn} = \sum_{l=1}^{17} v_l B_l \quad (4)$$

Para cada individuo n , el índice de bienes durables (IBD) se obtiene como una suma ponderada, donde v_l es el ponderador, B es una variable dicotómica con valor 1 si el individuo posee en su hogar

el bien l , y valor 0 si el individuo no posee el bien l . Para $l = 1, \dots, 17$, correspondiente a los bienes durables incluidos en la encuesta, detallados en el anexo.

En este trabajo se presentan tres alternativas de ponderación. Según los criterios de ponderación utilizados se obtuvieron los índices IBD1, IBD2 e IBD3, a saber:

- IBD_1 : Los 17 bienes reciben la misma ponderación ($v_l=1$, $l=1..17$). El índice de bienes durables en este caso captura la expresión del mayor capital del hogar evidenciado por la posesión simultánea de todos los electrodomésticos considerados.
- IBD_2 : Se clasificaron (por criterio) los bienes en dos grupos y se asignaron ponderaciones diferentes a los bienes de cada grupo. Un primer grupo (compuesto por horno a microondas, video reproductor; computadora; Internet, video filmadora y aire acondicionado) reunió a los bienes de mayor valor y asociados subjetivamente con un mejor pasar económico. El ponderador de los bienes de este conjunto asumió un valor 2, y el resto de los bienes (clasificados en un segundo grupo) mantuvo el valor 1 del ponderador.
- IBD_3 : A partir de la aplicación del Análisis Factorial de Correspondencias Múltiples (AFCM) al conjunto de datos con información completa, se procedió a analizar la aparición conjunta de determinados bienes en ciertos conjuntos de individuos, así como la ausencia simultánea de determinados bienes, en otros grupos. Este análisis (en el anexo se presentan los principales elementos considerados en este punto) junto con la verificación del porcentaje de hogares que poseía cada bien⁴, determinaron una segunda clasificación del conjunto total de bienes considerados. El primer grupo (con bienes que estaban presentes en un número más reducido de hogares, y al mismo tiempo, desde una perspectiva subjetiva corresponden a una manifestación de mayor poder adquisitivo) quedó conformado por los siguientes bienes: Computadora, Internet, Auto propio, TV por Cable, Aire acondicionado y Video Filmadora. La ponderación aplicada a la posesión de estos bienes fue de 2. Un segundo grupo (formado por bienes ampliamente difundidos, ausentes en gran parte de los hogares con muy reducida chance de poseer bienes) quedó conformado por: Ventilador, Microondas, Video reproductor, Secarropas, Equipo de música y Teléfono). El ponderador elegido para este subgrupo fue de 1,5. El resto de los bienes quedó compuesto por: Heladera, Heladera con freezer⁵, Lavarropas, Cocina a gas, Calefón y TV color. Este tercer subgrupo preservó un ponderado de valor 1 para la correspondiente posesión del bien.

La aplicación del AFCM se desarrolló con una finalidad exclusivamente exploratoria, tendiente a analizar la posibilidad de reducir la dimensión del problema inicial, agrupando individuos que presentaran posesión conjunta simultánea de distintos paquetes de bienes, así como la no posesión simultánea de determinada serie de bienes de capital en el hogar. Se obtuvo la representación en los principales ejes factoriales considerando como variables activas cada una de las 18 variables correspondientes a la posesión de bienes durables. Cada variable tenía dos categorías posibles: poseer el bien o no poseerlo. En el anexo se detallan los principales indicadores de la representación conjunta de los datos en el mapa factorial. Además, a través de la técnica de AFCM se procedió a detectar cuáles variables, permitían clasificar a los alumnos en grupos homogéneos internamente y heterogéneos entre sí (a través de la agrupación en clusters). La clasificación resultante permitiría obtener una aproximación a una estratificación social de los casos, si bien el objetivo principal, como se estableció previamente, no es asignar los individuos a clusters sino imputar el valor numérico asociado con la riqueza del hogar de origen.

De lo presentado anteriormente surge que el INSE dependerá en última instancia de la alternativa que se elija para construir el IBD y de las ponderaciones w_l que se adopten. Respecto de estas

⁴ En general, la posesión de determinados bienes de amplia difusión en la sociedad no operará como un elemento importante al diferenciar grados de nivel socio-económico (como se pretende en este trabajo), aunque la no posesión de esos mismos bienes puede operar como un indicador de una situación de precariedad extrema. Por esta razón, todos los bienes disponibles ingresaron al índice, pero en esta versión más sofisticada, algunos con un peso relativo menor.

⁵ Cabe aclarar que los bienes Heladera y Heladera con freezer, en la encuesta aparecen prácticamente como sustitutos, de manera que no se esperó en ninguno de los índices de bienes durables que los hogares alcanzaran un puntaje máximo de posesión de los 18 bienes, sino de 17 bienes (ya que o bien poseen heladera o heladera con freezer). Asimismo, en diferentes pruebas que se desarrollaron aplicando Análisis de Componentes Principales, la expresión en el primer componente de la variable poseer heladera (sin freezer) implicaba un coeficiente negativo, en tanto para todas las demás variables, ocurría lo opuesto (poseer cada uno de los bienes determinaba un aporte positivo). Pese a esta información, se preserva con esta estructura, pero constituye un elemento a explorar que podría aportar un elemento interesante de discriminación entre quienes poseen algún tipo de heladera.

últimas, un análisis exploratorio realizados con Componentes Principales llevó a la conclusión que la elección de $w_i=1/3$ para todo i es apropiada (ver anexo, apartado c).

5. Resultados

En esta sección presentamos los resultados obtenidos en la construcción del índice de nivel socio-económico bajo las distintas alternativas seleccionadas y que corresponden a la fórmula 1 de la sección anterior. Para ello se presenta primeramente el índice educación del padre, luego el índice de hacinamiento y finalmente se ofrecen dos variantes del índice de bienes durables. de hacinamiento y del índice de bienes durables⁶.

La tabla II presenta las principales medidas estadísticas del índice de educación del padre para el país y para las regiones consideradas.

Tabla II: Índice de Educación del Padre

IE	Observaciones	Media	Desviación Estándar	Mediana
Capital Federal	17,473	69.02	32.12	60
GBA	58,621	51.69	35.28	60
Pampeana	92,275	51.74	34.8	60
NOA	33,957	50.12	36.07	60
NEA	26,763	43.36	36.24	40
Cuyo	22,502	46.34	34.49	40
Patagonia	12,176	46.11	34.74	40
País	263,767	51.09	35.47	60

Fuente: Elaboración propia con datos de ONE 2000 (Base de Alumnos).

Las regiones que presentan valores comparativamente más altos resultados son Capital Federal, Gran Buenos Aires y Región Pampeana. En los tres casos, la media regional del índice de educación del padre es mayor a la media del país y su distribución es sesgada hacia la izquierda indicando una mayor concentración en valores altos del índice, mientras que en el resto de las regiones la distribución es sesgada hacia la derecha. Se destaca Capital Federal con una media 35.2% superior a la media nacional, 69 y 51 respectivamente. La región que arroja menor nivel educativo en los padres de los alumnos es la región Noreste. La media del índice de educación del padre es 43.36.

Con respecto a la condición de hacinamiento los resultados del indicador se exponen en la Tabla III. En este caso, sólo Capital Federal, Gran Buenos Aires y la Región Pampeana arrojaron medias superiores en el indicador de hacinamiento al promedio nacional. Sin embargo, el promedio de la Patagonia se encuentra levemente por debajo de la media del índice del total del país. El norte del país (NEA y NOA) se ubican en los valores más bajos del indicador.

Tabla III: Indicador de Hacinamiento

IH	Observaciones	Media	Desviación Estándar	Mediana
Capital Federal	17,473	22.48	12.16	19.51
GBA	58,621	17.39	10.79	15.12
Pampeana	92,275	18.11	11.53	15.12
NOA	33,957	15.58	10.16	12.2
NEA	26,763	14.93	9.66	12.2
Cuyo	22,502	16.05	10.08	14.29
Patagonia	12,176	16.90	10.78	15.12
País	263,767	17.36	11.05	15.12

Fuente: Elaboración propia con datos de ONE 2000 (Base de Alumnos).

⁶ Los cálculos vinculados al IBD_2 e $INSE_2$ pueden solicitarse a los autores.

Las siguientes tablas IV y V presentan dos versiones del índice de bienes durables. En la tabla IV los bienes fueron ponderados con igual peso mientras que en la tabla V se aplicó una ponderación diferenciada según se explica en la sección 4. b. anterior.

Tabla IV: Índice de Bienes Durables 1

IBD₁	Observaciones	Media	Desviación Estándar	Mediana
Capital Federal	17,473	76.65	16.19	76.47
GBA	58,621	66.43	17.80	64.71
Pampeana	92,275	64.66	18.01	64.71
NOA	33,957	52.98	22.42	52.94
NEA	26,763	50.14	23.95	52.94
Cuyo	22,502	60.06	18.48	58.82
Patagonia	12,176	61.60	18.19	64.7
País	263,767	62.34	20.36	64.71

Fuente: Elaboración propia con datos de ONE 2000 (Base de Alumnos).

Al analizar las diferencias entre regiones debidas al número de bienes durables poseídos por el hogar de origen (*IBD₁*), Capital Federal se ubica a una apreciable distancia por encima del promedio, así como las regiones del Norte del país (NEA y NOA), se ubican por debajo, destacándose en estas últimas regiones una heterogeneidad superior entre los individuos que en los demás casos.

El ranking determinado por los valores medios del indicador de posesión de bienes durables con ponderaciones diferenciadas para grupos de electrodomésticos que se presentan en la tabla V resulta similar al de ponderaciones simples expuestas en la tabla IV. Sin embargo, las relaciones entre las medias de las regiones se magnifica en el segundo caso (por ejemplo, con el *IBD₁*, la relación entre el promedio de Capital Federal, que detenta el máximo promedio, con el promedio de NEA, es de 1.53; en tanto para el *IBD₃*, este cociente es de 1.6).

Tabla V: Índice de Bienes Durables 3

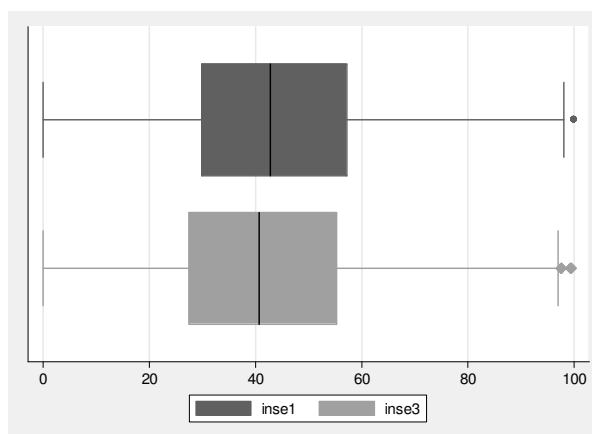
IBD₃	Observaciones	Media	Desviación Estándar	Mediana
Capital Federal	17,473	71.94	17.71	74.07
GBA	58,621	60.43	18.87	59.26
Pampeana	92,275	58.34	18.88	57.41
NOA	33,957	47.59	22.14	46.29
NEA	26,763	45.08	23.40	44.44
Cuyo	22,502	53.29	19.23	51.85
Patagonia	12,176	55.62	18.89	55.56
País	263,767	56.42	20.89	57.41

Fuente: Elaboración propia con datos de ONE 2000 (Base de Alumnos).

El índice de nivel socio-económico:

La figura I permite observar que el *INSE₁*, que incorpora el *IBD₁* presenta en general valores levemente más alto que el *INSE₃*, que incorpora el *IBD₃*. En cuanto a la dispersión (capturada a través de la desviación estándar) no se observan diferencias importantes.

Figura I: Diagrama de caja de $INSE_1$ e $INSE_3$

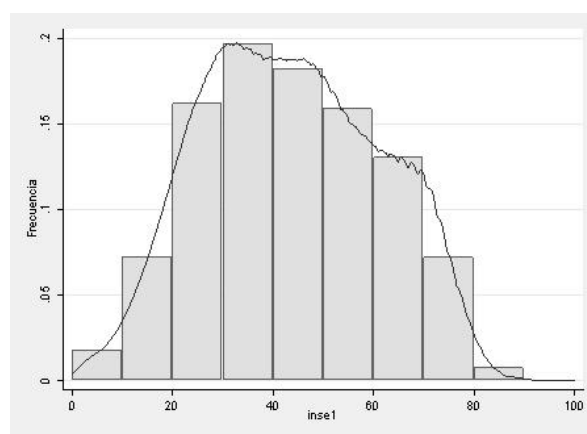


Fuente: Elaboración propia con datos de ONE 2000 (Base de Alumnos).

Los índices de nivel socio-económico obtenidos se distribuyen con un sesgo a la derecha (tal como se espera en general para distribuciones de ingresos o riquezas de hogares). Esta morfología evidencia la característica falta de equidistribución (con una alta participación de los hogares con baja acumulación, en este caso, del capital físico y humano).

De la comparación de las figuras II y IV surge por un lado que tanto en el caso de $INSE_1$ como el caso del $INSE_3$ la mayor concentración de las frecuencias ocurre entre el tercer y quinto decil, pero que la concentración de frecuencias es más alta cuando se utiliza el $INSE_3$.

Figura II: Distribución de Frecuencias del $INSE_1$



Fuente: Elaboración propia con datos de ONE 2000 (Base de Alumnos).

La tabla VI presenta una desagregación del $INSE_1$ por regiones.

Tabla VI: Índice de Nivel Socio-económico 1

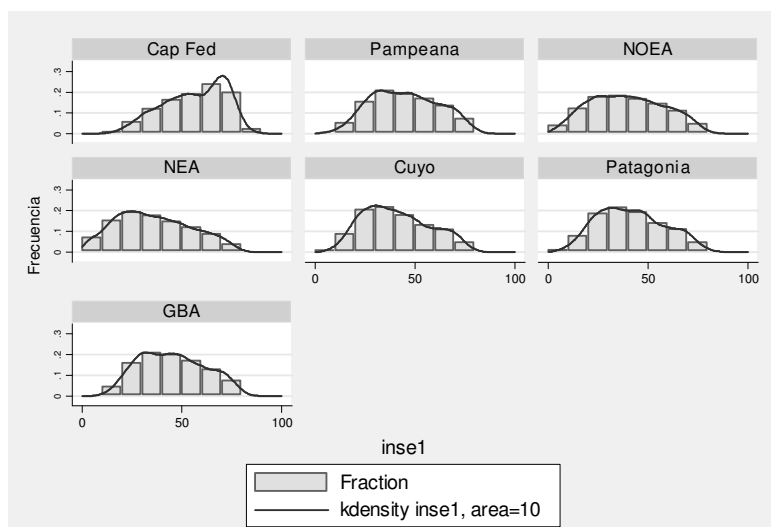
$INSE_1$	Obs.	Media	Desviación Estándar	Mediana
Cap. Fed.	17,473	56.05	15.82	57.98
GBA	58,621	45.17	16.59	44.15
Pampeana	92,275	44.84	16.95	43.87
NOA	33,957	39.56	18.39	38.76
NEA	26,763	36.14	18.85	34.35
Cuyo	22,502	40.81	16.91	38.97
Patagonia	12,176	41.54	16.95	40.14
País	263,767	43.60	17.76	42.80

Fuente: Elaboración propia con datos de ONE 2000 (Base de Alumnos).

El ordenamiento de las regiones de acuerdo con los promedios del indicador no contradicen las hipótesis vinculadas con el desarrollo relativo de las regiones así como de las características de pobreza relativa. El mayor promedio del $INSE_1$ se registra en Capital Federal, casi un 30% superior al promedio del país, seguido por el conurbano bonaerense y por la región Pampeana, con valores medios muy similares entre sí. Las regiones con promedios más bajos de $INSE_1$ (como también con mayor dispersión entre las observaciones de la variable), son NEA y NOA, un 17% y un 9%, ambas por debajo del total del país.

Como se observa en la Figura III que describe la distribución de frecuencias del $INSE_1$ por regiones, la situación de Capital Federal difiere sustantivamente del resto de las regiones reflejando la mayor riqueza relativa de su población.

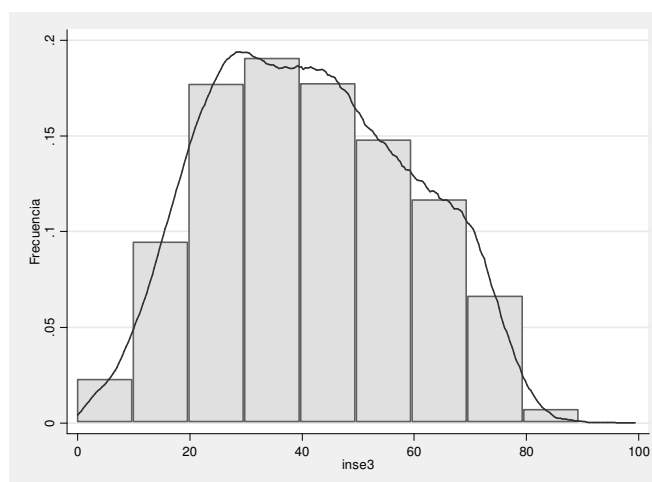
Figura III: Distribución de Frecuencias del $INSE_1$ por Regiones



Fuente: Elaboración propia con datos de ONE 2000 (Base de Alumnos).

En efecto, la distribución de los valores de $INSE_1$ para Capital Federal presenta un sesgo a la izquierda, lo cual implica que la mayor concentración de valores se registra para valores relativamente altos del índice. El intervalo con mayor cantidad de hogares es el correspondiente a entre 60 y 70 puntos, seguido por el de entre 70 y 80 puntos. En todas las demás regiones el índice presenta sesgo a la derecha siendo los casos más notorios el NEA y NOA.

Figura IV: Distribución de Frecuencias del $INSE_3$



Fuente: Elaboración propia con datos de ONE 2000 (Base de Alumnos).

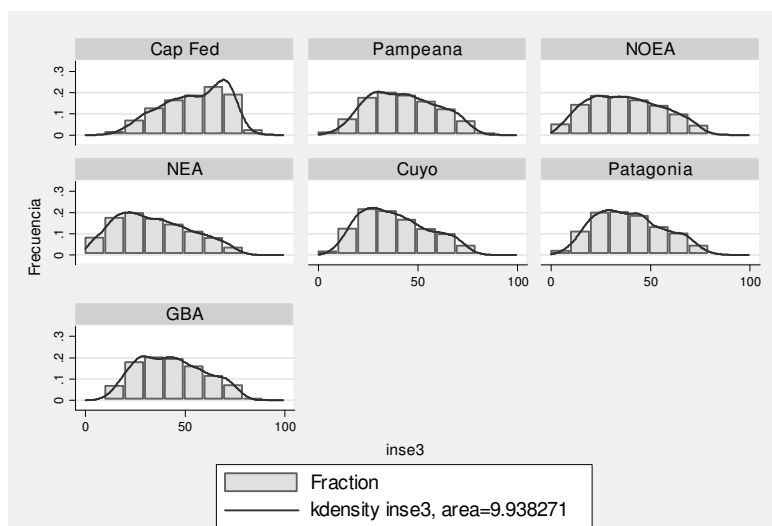
Tabla VIII: Índice de Nivel Socio-económico 3

INSE ₃	Obs.	Media	Desviación Estándar	Mediana
Capital Federal	17,473	54.48	16.23	56.53
GBA	58,621	43.17	16.89	42.01
Pampeana	92,275	42.73	17.21	41.71
NOA	33,957	37.76	18.34	36.64
NEA	26,763	34.46	18.74	32.34
Cuyo	22,502	38.56	17.14	36.53
Patagonia	12,176	39.55	17.17	38.10
País	263,767	41.62	17.95	40.70

Fuente: Elaboración propia con datos de ONE 2000 (Base de Alumnos).

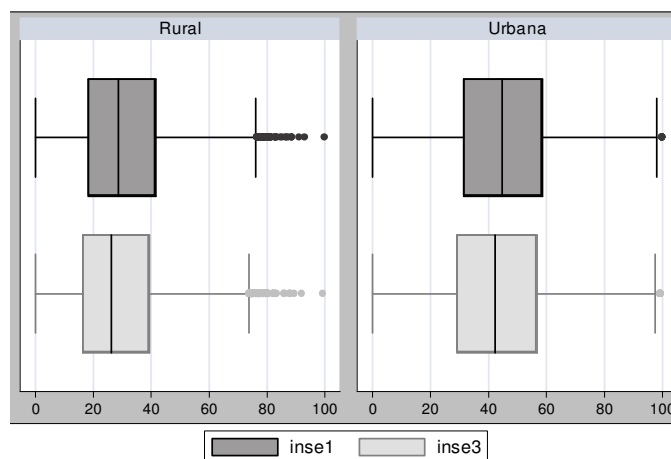
En el caso del $INSE_3$, se observa una mayor concentración de hogares (comparada con la del $INSE_1$) en valores bajos del indicador, y correspondientemente una menor acumulación en valores altos. Las diferencias entre regiones se conservan, tanto en cuanto a la forma de la distribución como al ranking surgido de los valores promedios.

Figura V: Distribución de Frecuencias del INSE₃ por Regiones



Fuente: Elaboración propia con datos de ONE 2000 (Base de Alumnos).

Figura VI: Diagrama de caja: INSE₁ e INSE₃ según área



Fuente: Elaboración propia con datos de ONE 2000 (Base de Alumnos).

Tal como puede verse en la Tabla IX que compara el $INSE_1$ y el $INSE_3$ según áreas rural y urbana

Tabla IX: $INSE_1$ e $INSE_3$ según área

Área	Obs	Media	Desviación Estándar	Mediana
$INSE_1$				
Rural	28,614	30.56	12.29	28.57
Urbana	235,119	45.18	17.27	44.63
Nc.	34	39.18	19.44	40.33
$INSE_3$				
Rural	28,614	28.67	16.01	26.27
Urbana	23,5119	43.20	17.52	42.32
Nc.	34	37.02	19.37	37.09

Fuente: Elaboración propia con datos de ONE 2000 (Base de Alumnos).

los valores de los indicadores, en relación con el área de residencia de los individuos bajo estudio, como se esperaba, resultan inferiores en general para los residentes en áreas rurales con respecto a los residentes en áreas urbanas, con ambos indicadores. Las distribuciones permiten observar para ambas opciones de cálculo del INSE (ver Figura VI), que para los menores que asisten a establecimientos en zonas rurales, prácticamente tres cuartas partes se ubica en valores de 40 puntos o menos del indicador, en tanto en las zonas urbanas, menos de la mitad de los casos obtuvo puntajes menores a 40 puntos.

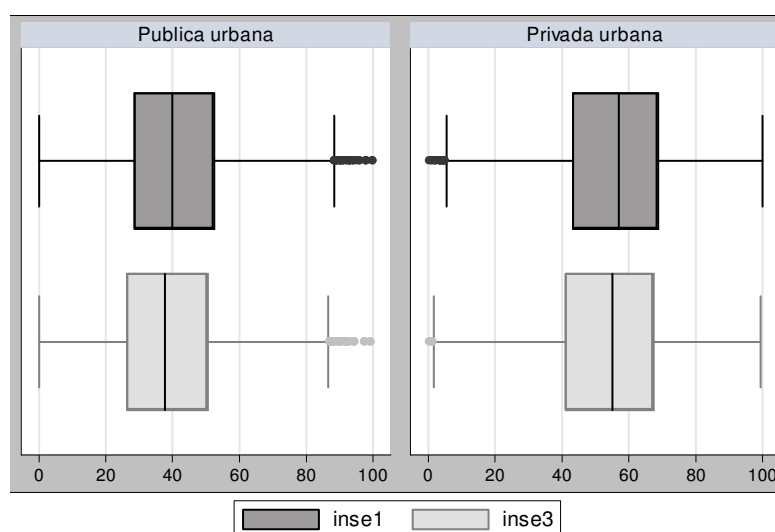
Como se observa en la Tabla X y en la figura VII, también en el caso del tipo de escuela al que residen los menores, los valores de los indicadores calculados son consistentes con las hipótesis previas. Los valores medios son superiores entre los individuos que asisten a escuelas privadas que los de individuos que asisten a escuelas públicas. Además, esta relación se observa no sólo en los valores promedio, sino también para el conjunto de la distribución: los valores muy bajos (tanto $INSE_1$ como $INSE_3$) son valores atípicos entre quienes asisten a escuelas privadas, y lo mismo ocurre con valores muy elevados para quienes concurren a escuelas públicas.

Tabla X: $INSE_1$ e $INSE_3$ según tipo de escuela

Área	Obs	Media	Desviación Estándar	Mediana
$INSE_1$				
Pública rural	28,614	30.56	12.29	28.57
Pública urbana	166,642	41.03	16.03	39.94
Privada urbana	68,477	55.29	15.96	57.00
Nc.	34	39.18	19.44	40.33
$INSE_3$				
Rural	28,614	28.67	16.01	26.27
Pública urbana	166,642	38.95	16.16	37.71
Privada urbana	68,477	53.53	16.38	55.05
Nc.	34	39.18	19.44	40.33

Fuente: Elaboración propia con datos de ONE 2000 (Base de Alumnos).

Figura VII: Diagrama de caja: INSE₁ e INSE₃ según tipo de escuela en el área urbana



Fuente: Elaboración propia con datos de ONE 2000 (Base de Alumnos).

6. Análisis de la validez externa de los resultados obtenidos

Como se mencionó previamente, además de generar un índice de nivel socio-económico a partir de la información disponible en las pruebas de calidad educativa, y analizar la consistencia de las caracterizaciones que se desprendían a partir de este índice para distintas subpoblaciones (regiones, áreas y tipo de escuelas), se desarrolló una prueba vinculada con la validez externa de una de las versiones de índices propuestos.

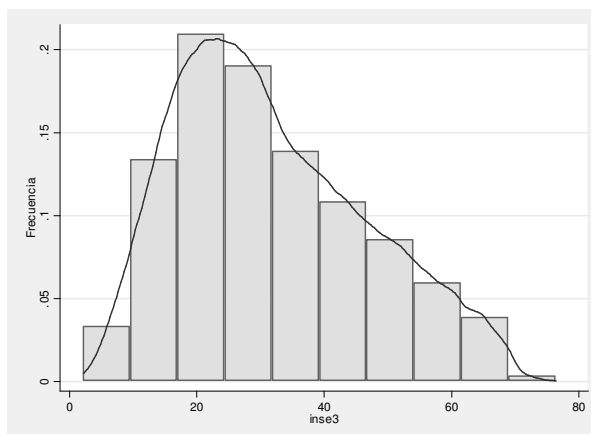
Para una primera validación del índice se trabajó con la última edición disponible de la Encuesta de Condiciones de Vida (ECV-Siempro 2001)⁷. La ECV-2001 se realizó en base a una muestra de 26.000 viviendas urbanas de todo el país. Se trata de un muestreo nacional urbano. El universo de la muestra abarca a la población residente en localidades de 5.000 o más habitantes. La misma representa aproximadamente el 96% de la población urbana del país, y el 84% de la población total. Las razones que motivan la utilización de una base como la descrita, se vinculan con las variables disponibles (variables que no están presentes en otros relevamientos periódicos como la Encuesta Permanente de Hogares).

La encuesta Siempro cuenta con información acerca de las tres dimensiones del INSE consideradas: educación (del jefe del hogar en este caso), hacinamiento (con un registro más detallado que en el caso de los cuestionarios complementarios a las pruebas educativas) y posesión de ciertos electrodomésticos. En este tercer aspecto cabe aclarar que el listado de bienes considerados no es exactamente el mismo, por lo que para la prueba que desarrollamos se consideraron sólo aquellos bienes indagados en el cuestionario de Siempro que también formaban parte de los indagados en las pruebas educativas⁸. Este aspecto (una limitación inherente a los datos disponibles) constituye una de las principales restricciones al ejercicio de validación externa desarrollado. Otro aspecto importante a tener en cuenta es que el relevamiento de la ECV se desarrolló durante el año 2001 (al año siguiente a la base más actualizada disponible hasta el momento de las pruebas educativas, utilizada para el análisis inicial). Precisamente durante el año 2001 se intensificó el proceso de empobrecimiento de algunas regiones (particularmente las más desfavorecidas, NEA y NOA, como también del Conurbano Bonaerense y la región Pampeana). Fue posible calcular el INSE para 19321 casos que contaban con información de todas las variables consideradas.

⁷ La encuesta de Condiciones de Vida (Siempro 2001) tiene como finalidad “proveer información que aporte a la caracterización actualizada de las condiciones de vida de la población urbana de la Argentina... (y) estimar el acceso diferencial a los mismos según el nivel de ingresos de la población” (Siempro, 2001, Objetivos y metodología, pp.2).

⁸ Los bienes acerca de los que se interroga en la encuesta Siempro son: cocina a gas, heladera sin freezer, heladera con freezer, lavarropas, teléfono, TV, TV por cable, video reproductor, PC, Internet, Automóvil (entre los que también son indagados en las pruebas educativas), calefactores, moto o ciclomotor, bicicleta (no indagados en las pruebas educativas). Los bienes acerca de los que sí se indaga en los cuestionarios de las pruebas pero no en la ECV son: calefón o termotanque, ventilador, aire acondicionado, microondas, secarropas, equipo de música y filmadora.

Figura VI: Distribución de Frecuencias del INSE 3- Base Siempro



Fuente: Elaboración propia con datos de ECV (Siempro, 2001).

La estructura de la distribución del $INSE_3$ para el total de datos presenta también un sesgo a la derecha (muestran una concentración elevada de casos en los niveles relativamente bajos del indicador y baja proporción con valores elevados) tal como se esperaba. Sin embargo, en relación con la distribución obtenida para los alumnos (ONE 2000), se observa una mayor concentración en valores bajos. Por ejemplo, en el caso de la muestra de Siempro, casi un 42% de los hogares alcanza un valor inferior a 30 puntos del $INSE_3$, en tanto entre los individuos de las pruebas educativas ese conjunto representa el 30% de los casos.

A partir de la aplicación análoga de los índices presentados en la sección 4, se obtuvieron para cada hogar presente en la base Siempro, indicadores del nivel socio-económico. En la Tabla XI presentan los valores promedios para la muestra completa como para las respectivas regiones.

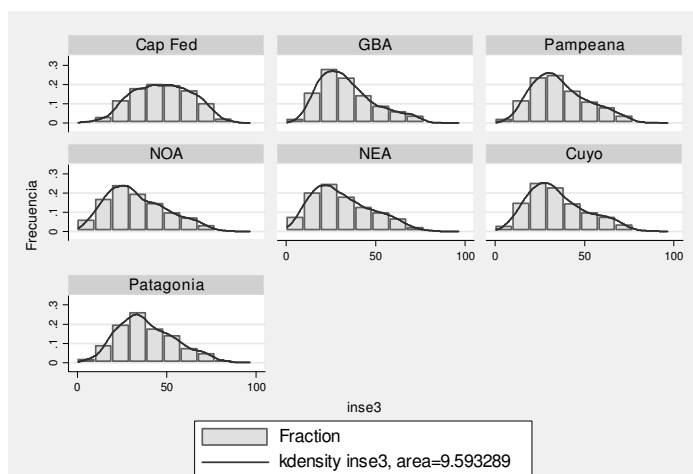
A diferencia del ranking registrado en el caso de los cuestionarios complementarios a las pruebas educativas, la segunda región en cuanto al INSE resulta la Patagonia (ambas por encima del promedio nacional). Las regiones con valores más bajos continúan siendo el NEA y el NOA, pero ahora apenas superadas por GBA.

Tabla XI: Índice de Nivel Socio-económico 3- Base Siempro

$INSE_3$	Obs.	Media	Desviación Estándar	Mediana
Cap. Fed	612	47.63	16.02	47.52
GBA	1,545	33.96	15.44	30.97
Pampeana	6,360	36.65	15.77	34.17
NOA	3,716	33.43	16.92	30.52
NEA	2,362	31.45	16.63	28.58
Cuyo	1,827	35.03	16.43	32.02
Patagonia	2,890	38.91	16.12	36.62
Total	19321	35.7	16.5	33.3

Fuente: Elaboración propia con datos de ECV (Siempro, 2001).

Figura VII: Distribución de Frecuencias del INSE₃ por Regiones- Base Siempre



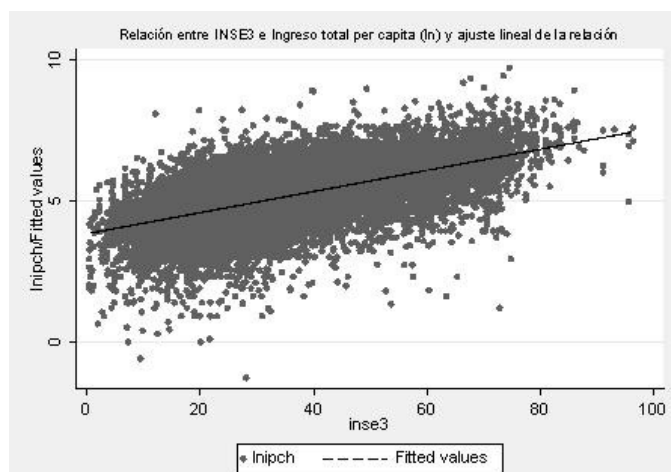
Fuente: Elaboración propia con datos de ECV (Siempre, 2001).

Dado que en el estudio Siempre se indagó acerca del ingreso familiar, resulta posible utilizar una medida derivada de éste, el ingreso familiar per capita, considerado en general una medida de la capacidad de compra del hogar más adecuada que el ingreso total, para analizar su correlación con la medida de INSE calculada para los mismos hogares.

Si bien es de gran interés este tipo de análisis para evaluar la calidad del índice generado, se espera que la relación lineal entre ingreso y las diferentes manifestaciones de capital del hogar no sea elevada en valor absoluto, pero sí que esta relación tendrá signo positivo. Si bien puede suponerse que la capacidad de obtener ingresos se correlaciona positivamente con el nivel educativo y con el mayor status social –en cuanto a favorecer el acceso a información, a conexiones, etc.–, esta relación tiene una gran variabilidad en términos del ingreso obtenido para cada nivel específico de capital acumulado. Por otra parte, cabe destacar que la información sobre ingresos del hogar no es conceptualmente equivalente a la noción de nivel socio-económico objeto de este estudio. El ingreso constituye una variable que se espera que esté positivamente relacionada con el nivel socio-económico por las razones expuestas, no obstante, el primero está más expuesto a oscilaciones coyunturales (ingresos extraordinariamente altos o bajos en un período, pero que no reflejen claramente la situación habitual de ingresos)⁹. En particular, con los datos de Siempre, se obtiene una correlación lineal de 0,59 entre el $INSE_3$ y el ingreso familiar per capita (en logaritmo natural)¹⁰, y las características de la asociación entre ambas variables pueden apreciarse en la figura VIII.

⁹ Además es importante considerar la diferencia entre ambos conceptos, destacando que la variable ingreso constituye una medida de flujo (cantidad percibida por unidad de tiempo), en tanto el nivel socio-económico, reflejo de una serie de acumulaciones del hogar en el tiempo, una variable stock.

Figura VIII: Relación entre INSE₃ e Ingreso familiar per capita



Fuente: Elaboración propia con datos de ECV (Siempro, 2001).

Finalmente, se desarrolló un ejercicio, cuyos resultados específicos para el conjunto de combinaciones consideradas puede observarse en la última sección del anexo, modificando en la construcción del INSE la ponderación de cada uno de los tres componentes (IE, IH y IBD). En los casos revisados que tratan de captar situaciones variadas para las ponderaciones de cada uno de los componentes, se observa que la correlación entre la medida de INSE respectivo generado con el logaritmo del ingreso per capita no difiere substancialmente, excepto para los casos extremos donde uno de los componentes tienen una ponderación muy reducida. Sin embargo, el aspecto del peso de los componentes en la especificación puede ser revisada a través de la aplicación de métodos econométricos si fuera de esperar una mayor relación entre el INSE obtenido y el ingreso del hogar.

7. Síntesis de resultados y conclusiones

Los principales aspectos vinculados con el cálculo de un índice de nivel socio-económico con información resultante de un estudio ya desarrollado se vinculan con el desarrollo de un indicador que sea satisfactorio al menos en tres aspectos para ser considerado válido:

- 1.- Alineación del constructo con el concepto teórico
- 2.- Capacidad del indicador de reflejar relaciones conocidas entre subconjuntos de datos
- 3.- Utilidad del indicador para reflejar la realidad que trata de captar en el contexto de conjuntos de datos diferentes al que le provee la información para su generación.

El primer aspecto fue considerado en el trabajo desde la generación de los índices mediante la revisión conceptual de los fenómenos reflejados en cada variable disponible y exploración de alternativas de medición de diferentes manifestaciones del capital del hogar. Esta consideración implicó en ciertos casos la exclusión de ciertas variables del análisis y la generación de ciertas variables a partir de la información inicial.

En este punto la principal limitación se asocia con la imposibilidad de contar con variables vinculadas al nivel socio-económico por no haber sido incorporadas al cuestionario. La principal ausencia (en relación con otros estudios sobre INSE, en algunos de los cuales se desarrollaron encuestas a niños como en el caso del ONE) es la variable ocupación del principal sostén o jefe de hogar.

El segundo aspecto estudiado se vincula con las relaciones entre subconjuntos de la población analizada en cuanto al constructo propuesto. Para el conjunto completo de hogares de alumnos en las pruebas, las características de la distribución de los datos en función de los puntajes asignados se condicen con lo esperado. En general, se verifican las relaciones previstas para la situación de las diferentes regiones del país en cuanto a su posicionamiento relativo en el puntaje de INSE obtenido. En los diferentes índices, los valores más elevados son registrados por los hogares de los alumnos de

¹⁰ La aplicación de una regresión (por MCO) con el ingreso per capita (en logaritmos naturales) como variable dependiente y el INSE como variable independiente arroja una estimación de la pendiente con signo positivo y significativamente distinta de cero (graficada en la figura VIII). El coeficiente de determinación R^2 , arroja un valor de 0.35.

Capital Federal, y los más bajos por las regiones del norte del país (NEA y NOA). También para otros subconjuntos, como área de residencia –urbano y rural- o tipo de establecimiento –público rural, público urbano y privado urbano, se registran situaciones en los términos esperados. Los valores de INSE son en general superiores para los hogares de los alumnos residentes en áreas rurales que para los residentes en zonas urbanas; y a su vez, superior para los menores que asisten a escuelas privadas que para los que asisten a escuelas públicas (entre quienes residen en áreas urbanas).

El tercer aspecto considerado, de mayor dificultad práctica, en función de la información necesaria, y del período de referencia, fue revisado a partir de la consideración de los datos de la Encuesta de Condiciones de Vida (Siempre 2001). La aplicación del INSE, en la versión 3, arroja una distribución apropiada de los datos, con una mayor proporción de casos con puntajes relativamente bajos que la surgida de las pruebas ONE. Asimismo, si bien la situación de los residentes en las diferentes regiones resulta aparentemente adecuada en cuanto a las situaciones extremas (Capital Federal en el tope y las regiones del norte con los mínimos valores), el ordenamiento de las restantes regiones difiere del correspondiente a los datos de alumnos para el año 2000. La asociación entre la variable asignadora de puntajes de acuerdo al capital del hogar (INSE3) y el ingreso del hogar (en términos per capita y en logaritmo natural) resulta dentro de los valores esperados.

Finalmente, y vinculado con la posibilidad de optimizar el cálculo de indicadores del NSE, se propone la inclusión de alguna pregunta relacionada con la ocupación de los padres del menor en el cuestionario complementario de las pruebas educativas. También se recomienda que a formulación de las preguntas debe ser acorde a la capacidad de los alumnos de comprender y registrar una respuesta lo más próxima a la situación indagada.

8. Referencias

Álvarez, G., Lucarini, A., y Mario, S. (2001). “La pobreza a partir de los datos censales: nuevos desarrollos basados en la Capacidad Económica de los hogares.”. En *Anales de la VII Jornadas Argentinas de Estudios de Población*. Neuquén. 2001

Asociación Argentina de Marketing (2002). Presentación del Nuevo índice de nivel socio-económico.

Braun y Mirlevy (2005). “Why Our Institutions about Testing Don’t Work?” *Phi Delta Kappan The Professional Journal of Education*, March 2005.

Caro, D. (2002). “Estimación del Nivel socio-económico de las familias: Propuesta metodológica para la Evaluación de Rendimiento del 2001”. Unidad de Medición de la Calidad Educativa. Ministerio de Educación de Perú.

CIPPEC (2004). “Los Estados provinciales frente a las brechas socio-educativas. Una sociología política de las desigualdades educativas en las provincias argentinas”. En <http://www.cippec.org/espanol/educacion/archivos/Informe%20final%20CIPPEC%20-%20PREAL.pdf>.

Clark, L. A. y Watson, D. (1995). « Constructing Validity : Basic Issues in Objective Scale Development ». *Psychological Assessment*, 1995, Vol. 7, No. 3, 309-319.

Coudouel, A., Hentschel, J., Wodum, Q. Medición y análisis de la pobreza. Técnicas básicas y problemas interrelacionados. The World Bank. Washington. 2002.

Davis, P., Jenkin, G., Coope, P. (2002). New Zealand Socio-economic Index 1996. An update and revision of the New Zealand Socio-economic Index of Occupational Status. Research Report No. 20. Statistics New Zealand.

Gómez, A., Mario, S. y Olmos, F. “Índice de Privación Material de los Hogares (IPMH): desarrollo y aplicación con datos del Censo 2001”. En *Anales de la VII Jornadas Argentinas de Estudios de Población*. Taí del Valle. 2003

INDEC. Acerca del método utilizado para la medición de la pobreza en Argentina. Documento preparado por la Dirección Nacional de Encuestas de Hogares. INDEC. Buenos Aires. 2003.

Lebart, L., Morineau, A. y Piron, M. (1997). *Statistique exploratoire multidimensionnelle*. París, Dunod.

Lee, H. “Poverty Measurement in Argentina” en *Argentina – Crisis and Poverty 2003, A Poverty Assessment*. Background Paper Number 8. The World Bank. Washington. 2003.

Levin H. (1971) “A New Model of School Effectiveness” en Mood, A. (comp.) *Do Teachers Make a Difference?* Report to the Secretary of HEW. Washington, DC, US Printing Office.

Llach, J., Montoya, S. y Roldán, F. (1999). *Educación para todos*. Córdoba.

Marks, G. (1999) “The Measurement of Socioeconomic Status and Social Class in the Longitudinal Surveys of Australian Youth Project”. Australian Council for Educational Research. Technical Paper No. 14. November 1999.

- Minujín, A. "Transformaciones en la estructura social argentina". En Minujín, A., Beccaria, L. y otros. *Cuesta abajo. Los nuevos pobres: efectos de la crisis en la sociedad argentina*. UNICEF-Lozada. Buenos Aires. 1992.
- Monk, D. (1989). "The Education Production Function: Its Evolving Role in Policy Analysis". *Educational Evaluation and Policy Analysis*. Spring 1989, Vol. 11, No. 1.
- Peña, D. (2002) *Análisis de datos multivariantes*. McGraw Hill. Madrid
- Perry, B. (2002). "The Mismatch between Income Measures and Direct Outcome Measures of Poverty". *Social Policy Journal of New Zealand* - Issue 19.
- Rowe, K. and Hill, P. (1995). "Methodological Issues in Educational Performance and School Effectiveness Research: A Discussion with Worked Examples". *Australian Journal of Education*, Vol. 39, No. 3, 1995.
- SAIMO. (2004). Presentación final de revisión INSE 2002. Sociedad Argentina de Investigación de mercados y opinión pública. En www.saimo.org.ar/archivos/presentacionfinal.zip.
- Siempre (2001). Encuesta de Condiciones de Vida 2001. Objetivos y Metodologías. Ministerio de Acción Social.
- Taccari, D. y Baruzzi, G. Medición de la Vulnerabilidad Social de la Población Escolar y su Participación Relativa en los Establecimientos Educativos: Una Propuesta Metodológica. Informe Metodológico N°2. Diniece.
- UNESCO (2000). Primer estudio internacional comparativo sobre Lenguaje, Matemática y factores asociados, para alumnos de tercer y cuarto grado de la educación Básica.
- Viguer Seguí, P. y Serra Desfilis, E. (1996) "Nivel socio-económico y calidad del entorno familiar en la infancia". *Anales de Psicología*, 1996, 12(2), 197-205.

9. Anexo

a. Información disponible sobre las variables seleccionadas para la construcción de los componentes del INSE

En la tabla XII se detallan las variables disponibles y las preguntas utilizadas en los cuestionarios complementarios a las pruebas del ONE 2000.

Tabla A1. Variables disponibles

Variable	Pregunta a los alumnos
Cantidad de miembros del hogar	¿Cuántas personas viven permanentemente en tu casa, contándote también a tí?
Número de habitaciones	¿Cuántas habitaciones tiene la vivienda en que vives, sin contar la cocina y el baño?
Agua corriente por red	¿Dentro de tu casa tienen agua por cañería?
Instalación eléctrica	Tu casa, ¿tiene instalación eléctrica?
Posesión de cada bien	¿Podrías indicar las cosas que SÍ hay y las cosas que NO hay en tu casa?
	Calefón o termotanque
	Heladera
	Freezer o heladera con freezer
	Cocina a gas
	Ventilador
	Horno microondas
	Video cassette
	Lavarropas
	Secarropas
	Computadora
	Internet
	Equipo de música
	Automóvil propio
	Teléfono
	TV color
Tv por cable	
Video filmadora	
Aire acondicionado	
Nivel educativo del padre	¿Hasta qué nivel educativo llegó tu papá?
Nivel educativo de la madre	¿Hasta qué nivel educativo llegó tu mamá?

En total, hay 270.430 casos (44% del total) en la base de datos de Alumnos en los cuales la respuesta sobre nivel de instrucción del padre no está registrada. De éstos, el 80% (216.600 casos) tiene respuestas en otras variables del Cuadernillo. Esta pérdida de respuestas puede deberse a la formulación de la pregunta, lo cual constituye un aspecto a revisar en los Operativos, ya que en el cuestionario de 1999¹¹, donde la pregunta parece ser más simple, la tasa de respuesta es del 84.4%.

Tabla A2. Nivel educativo del padre por regiones

Región	Casos	Primario incompleto	Primario completo	Secundario incompleto	Secundario completo	Univ. o Terciario incompleto	Univ. o Terciario completo
Cap. Fed	17,473	5.00	10.56	12.05	21.98	7.53	42.88
GBA	58,621	13.49	22.65	11.66	20.60	7.34	24.27
Pampeana	92,275	13.08	21.94	13.00	20.80	7.56	23.62
NOA	33,957	16.54	21.65	11.70	18.51	7.98	23.62
NEA	26,763	22.96	24.37	10.41	16.59	6.52	19.15
Cuyo	22,502	16.49	24.86	13.96	18.62	7.29	18.78
Patagonia	12,176	18.32	21.97	14.94	19.24	6.62	18.91
País	263,767	14.61	21.80	12.38	19.85	7.39	23.96

Fuente: Elaboración propia con datos de ONE 2000 (Base de Alumnos).

Se presentan en la tabla XIV los porcentajes de hogares con y sin hacinamiento (se considera que un hogar enfrenta la situación de hacinamiento si el indicador personas por habitación es mayor que 2 y dentro de esta categoría se distingue la situación como crítica cuando la relación excede a 3). En la tabla XV, se indican los porcentajes de menores que refirieron que poseían cada bien para el total del país y para cada una de las regiones.

¹¹ Pregunta ONE 1999: Con respecto a la educación de tu papá (Marca en cada pregunta SÍ o NO, según corresponda). ¿Tu papá...1 sabe leer y escribir? / 2 fue a la escuela primaria? / 3 terminó la escuela primaria? / 4 fue a la escuela secundaria? / 5 terminó la escuela secundaria? / 6 fue a la Universidad? / 7 terminó la Universidad?.

Tabla A3. Condición de hacinamiento por regiones

Región	Casos	Con hacinamiento		Sin hacinamiento
		Crítico	No crítico	
Cap. Fed	17,473	4.38	9.25	86.37
GBA	58,621	10.40	19.38	70.22
Pampeana	92,275	9.68	19.16	71.16
NOA	33,957	14.93	22.66	62.42
NEA	26,763	16.13	23.30	60.56
Cuyo	22,502	11.95	22.34	65.71
Patagonia	12,176	10.55	20.13	69.32
País	263,767	11.05	19.74	69.21

Fuente: Elaboración propia con datos de ONE 2000 (Base de Alumnos).

Tabla A4. Porcentaje de hogares que posee cada bien por región

	Cap. Federal	GBA	Pampeana	NOA	NEA	Cuyo	Patagonia	Total
Calefón o termotanque	92.82	83	87.01	62.8	50.8	81.97	84.04	79.1
Heladera	7.02	12.9	18.37	18	16.9	22.86	15.57	16.5
Heladera con freezer	92.44	85.8	79.28	73.4	70.4	73.68	80.63	79.5
Cocina a gas	97.36	94.4	96.84	92.1	91.1	96.6	97.13	95.1
Ventilador	90.15	87.8	83.5	73.2	83.3	77.39	40.37	81
Horno a microondas	66.81	44.7	35.43	26.6	23.2	28.9	30.42	36.4
Videoreproductor	85.57	67.7	55.73	42.4	37	48.18	54.03	56
Lavarropas	93.24	93.5	93.15	74.6	75.7	93.26	93.43	89.1
Secarropas	59.54	74.6	72.61	49.4	42.9	64.95	68.68	65.4
PC	68.96	38.6	33.79	20	16.1	26.16	35.31	33
Internet	43	19.1	13.35	8.77	6.96	10.27	15.24	15.2
Equipo de música	93.92	88.6	83.53	78.2	76.9	79.75	85.98	83.8
Automóvil	60.36	52.1	62.76	39.6	39.9	62.96	64.53	55
Teléfono	93.13	77.9	74.62	50.1	44.3	69.22	74.98	69.9
Tv color	99.03	97.9	96.18	91.4	82.6	95.73	95.8	94.7
Tv por cable	84.5	67.1	75.58	69.2	58.4	60.18	75.45	70.4
Videofilmadora	40.49	23.9	17.66	12.5	11.1	14.16	18.74	19
Aire acondicionado	34.66	19.8	19.93	18.4	24.7	14.72	16.91	20.6

Fuente: Elaboración propia con datos de ONE 2000 (Base de Alumnos).

b. Aplicación de AFCM al conjunto de bienes

Se incorporaron como variables activas del análisis la totalidad de bienes y servicios incluidos en el cuestionario (con las modalidades respectivas, poseer o no poseer cada uno de los bienes). El procesamiento estadístico se desarrolló utilizando SPADN. A continuación se presentan los principales resultados (coordinadas en los primeros 5 ejes factoriales y valores test de cada una de las modalidades en los ejes; el plano factorial).

Tabla A5. Histograma de los 5 primeros valores propios

Número	Valor propio	% acum..
1	0.2581	25.81
2	0.0992	35.73
3	0.0869	44.41
4	0.0512	49.54
5	0.0494	54.48

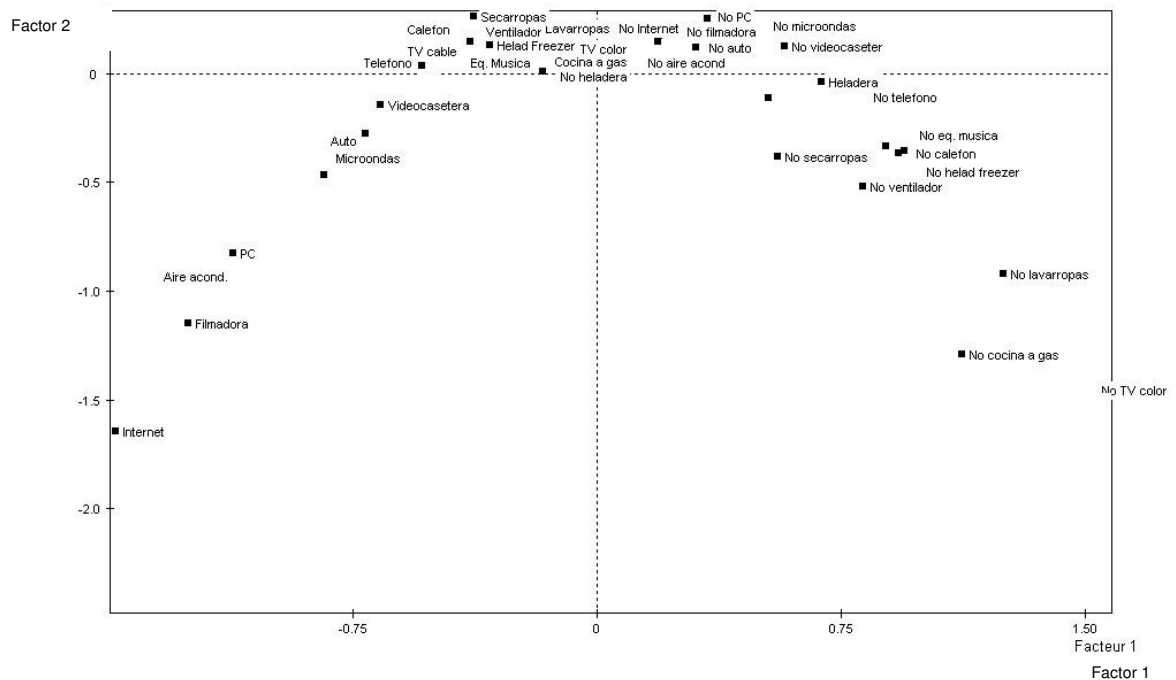
Fuente: Elaboración propia (Base alumnos)

Tabla A6. Coordenadas y valores test de las modalidades (Ejes 1 a 5)

Variable	Valor test					Coordenadas				
	1	2	3	4	5	1	2	3	4	5
Modalidad (posee)										
Calefón	99.9	99.9	71.0	70.9	-67.9	-0.39	0.15	0.09	0.09	0.08
Heladera	99.9	-9.6	99.9	-13.1	9.8	0.69	-0.04	1.85	-0.05	0.04
Heladera con freezer	-99.9	99.9	-99.9	-12.8	-18.2	0.33	0.13	-0.45	-0.01	-0.02
Cocina a gas	-99.9	99.9	59.2	-99.9	-99.9	0.08	0.09	0.03	0.18	-0.10
Ventilador	-99.9	99.9	52.7	-62.8	73.5	0.26	0.16	0.06	0.07	0.08
Microondas	-99.9	-99.9	-6.2	-46.5	47.2	0.84	-0.47	0.02	0.14	0.15
Videocassetera	-99.9	-68.9	39.8	49.0	33.9	0.67	-0.14	0.08	0.10	0.07
Lavarropas	-99.9	99.9	82.3	80.8	-41.4	0.24	0.18	0.07	0.07	-0.04
Secarropas	-99.9	99.9	71.7	99.9	23.1	0.38	0.26	0.12	0.38	0.04
PC	-99.9	-99.9	43.2	-27.9	-82.4	-1.12	-0.83	0.15	-0.10	-0.29
Internet	-99.9	-99.9	29.0	-77.9	-45.7	1.48	-1.65	0.17	0.46	-0.27
Equipo de música	-99.9	88.8	31.3	-38.3	99.9	0.24	0.09	0.03	0.04	0.16
Auto	-99.9	-93.2	39.1	99.9	-99.9	0.71	-0.28	0.12	0.33	-0.80
Teléfono	-99.9	20.6	60.5	99.9	-61.3	0.54	0.03	0.10	0.19	-0.10
Tv color	-99.9	99.9	77.0	-99.9	25.9	0.15	0.14	0.05	0.07	0.02
TV cable	-99.9	53.2	45.4	-53.3	99.9	-0.45	0.10	0.08	-0.10	0.39
Filmadora	-99.9	-99.9	20.3	-36.8	71.6	-1.26	-1.15	0.10	-0.18	0.36
Aire acondicionado	-99.9	-99.9	18.9	-45.2	99.9	-1.08	-0.85	0.09	-0.21	0.60

Fuente: Elaboración propia (Base alumnos)

Figura A1. Plano factorial (factores 1 y 2)



c. Aplicación de ACP a los índices parciales

Con la finalidad de analizar en qué medida la conjunción de los tres componentes del índice de nivel socio-económico explicaban la variabilidad total asociada con las manifestaciones estudiadas de capital del hogar, se aplicó la técnica de análisis de componentes principales a la muestra completa de trabajo. Los resultados indican que el primer componente capta aproximadamente el 59% de la variabilidad conjunta, y la expresión en el primer eje (utilizada como aproximación de la posible ponderación de los tres componentes) brinda valores similares para cada uno de ellos, con una leve mayor participación del componente de bienes durables.

Las variables activas fueron los respectivos índices de educación del padre, índice de hacinamiento e índice de bienes durables, donde los tres corresponden a variables cuantitativas, con escalas iguales de 0 a 100.

Tabla A7. Aplicación de componentes principales a componentes de INSE3

Componente	Valor propio	Diferencia	Proporción	Acumulada
1	1.77613	1.06070	0.5920	0.5920
2	0.71543	0.20699	0.2385	0.8305
3	0.50844	0	0.1695	1.0000

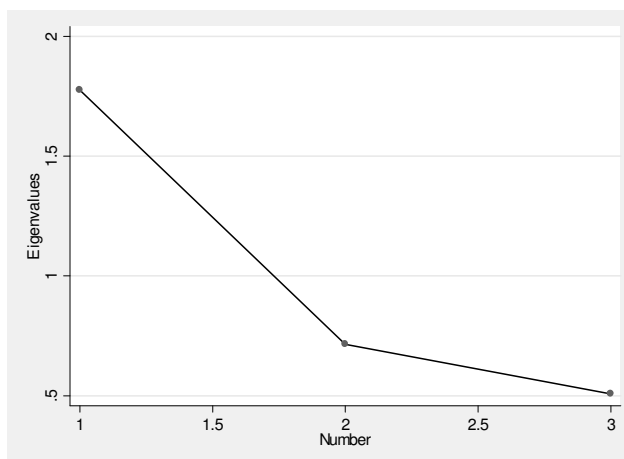
Fuente: Elaboración propia (Base alumnos)

Tabla A8. Coordenadas de los subíndices en el primer componente

Sub índice	Coordenadas
Índice de educación del padre	0.559
Índice de hacinamiento	0.547
Índice de bienes durables	0.622

Fuente: Elaboración propia (Base alumnos)

Figura A2. Importancia del primer valor propio



d. Evaluación de estructuras alternativas de ponderación de los componentes del INSE

Con los datos de la base de Siempre se desarrolló una prueba a partir de la utilización de una serie de distintos ponderadores de los tres componentes del INSE.

Los resultados obtenidos (en cuanto al INSE₃ promedio para la muestra total como al grado de asociación lineal del indicador generado y el indicador de ingreso –logaritmo natural del ingreso total familiar per capita-) se detallan en la siguiente tabla.

Tabla A9. Prueba de alternativas de ponderación de los componentes del INSE₃

IE	IH	IDB3	Promedio de INSE ₃	Coefficiente de correlación lineal con Ingreso familiar*
1/3	1/3	1/3	35.7	0,5900
0,6	0	0,4	43.9	0,4934
0	0	1	51.3	0,4879
0,2	0	0,8	48.8	0,5263
0,4	0	0,6	46.3	0,5259
0,8	0	0,2	41.4	0,4474
1	0	0	38.9	0,4013
0	0,2	0,8	44.8	0,5438
0,2	0,2	0,6	42.4	0,5845
0,4	0,2	0,4	39.9	0,5636
0,6	0,2	0,2	37.4	0,5056
0,8	0,2	0	34.9	0,4424
0	0,4	0,6	38.4	0,6005
0,2	0,4	0,4	35.9	0,6336
0,4	0,4	0,2	33.4	0,5788
0,6	0,4	0	30.9	0,4948
0	0,6	0,4	31.9	0,6219
0,2	0,6	0,2	29.4	0,6306
0,4	0,6	0	26.9	0,5467
0	0,8	0,2	25.4	0,5579
0,2	0,8	0	22.9	0,5437
0	1	0	18.9	0,4294

Fuente: Elaboración propia con datos de Siempro 2001. *Ingreso total familiar per capita (en logaritmo natural)

En cuanto a la correlación con el ingreso, en los casos simulados el signo resultante no difiere del esperado, y la asociación lineal resulta más baja cuando uno solo de los componentes es considerado como indicador del NSE. En cuanto al promedio, como puede observarse, las diferentes estructuras de ponderación producen cambios sustanciales en la medida de tendencia central. Aspectos vinculados con la estructura de ponderación de los componentes serán todavía tema de análisis.