



Universidad Nacional de La Plata



La Cobertura en Salud en el Área Metropolitana. Un Análisis Empírico en Base a Modelos de Elección Binaria y Multinomial

Mariana Marchionni

Documento de Trabajo Nro. 12
Abril 1999

Trabajo de Tesis de la Maestría en Economía,
Facultad de Ciencias Económicas, Universidad Nacional de La Plata.

**LA COBERTURA EN SALUD EN EL ÁREA METROPOLITANA.
UN ANÁLISIS EMPÍRICO EN BASE A MODELOS DE ELECCIÓN
BINARIA Y MULTINOMIAL.**

Por Mariana Marchionni

Director de tesis: Walter Sosa Escudero

La Plata, noviembre de 1998

RESUMEN

El objetivo central de este trabajo es explorar la relación entre las características socioeconómicas y demográficas de los individuos y la cobertura en salud. En base a la Encuesta Permanente de Hogares de 1992 para el Área Metropolitana de Buenos Aires, se hace un análisis empírico tanto de la determinación de la situación de cobertura (tener o no un seguro de salud) como de la forma de cobertura (el tipo de seguro). Se estima la probabilidad de tener cobertura mediante un modelo logit, mientras que un modelo logístico multinomial se utiliza para modelar la probabilidad de las distintas formas de seguro. Como resultado de este ejercicio se encuentra que las características socioeconómicas de los individuos afectan tanto la probabilidad de tener cobertura como la del tipo de cobertura. El artículo presenta un detallado análisis cualitativo y cuantitativo de los hallazgos. Como una contribución adicional, se propone una metodología que permite predecir la cobertura de la población a partir de información limitada usualmente disponible.

ABSTRACT

The primary objective of the present paper is to explore the relationship between health coverage and the socio-economic and demographic characteristics at the individual level. The empirical analysis, based on the Permanent Household Survey of 1992 for the Metropolitan area of Buenos Aires, is directed to explore not only the determinants of coverage (whether to have health insurance or not) but also the type of insurance. The probability of being covered is estimated with a Logit Model, while for the estimation of the probability of having each different kind of insurance a Multinomial Logit Model is used. The results of this exercise reveal that the socio-economic characteristics of the individuals affect both the probability of being covered and the probability of having each type of health insurance. The article presents a detailed qualitative and quantitative analysis of these findings. Moreover, a methodology to predict the population coverage using limited information is presented as an additional contribution.

La cobertura en salud en el Área Metropolitana. Un análisis empírico en base a modelos de elección binaria y multinomial.

por Mariana Marchionni*

1. Introducción

La cobertura en salud juega un papel decisivo en la determinación del nivel de utilización de servicios de salud. De ahí que su análisis resulte relevante, en especial con miras al diseño de políticas en el plano de la Salud Pública.

La incertidumbre acerca de futuras situaciones de enfermedad es lo que motiva una demanda por un seguro de salud. El costo de la atención médica que requiere un enfermo es lo que hace que el riesgo de enfermarse se traduzca en un riesgo financiero, contra el cual es posible asegurarse.

La ventaja que tienen los seguros es que reducen el precio efectivo pagado por los consumidores asegurados en el momento en que compran un bien o servicio. En particular, el seguro de salud reduce los precios percibidos por el paciente en el momento en que contrata los servicios médicos. La respuesta a estos menores precios es una mayor demanda de servicios de salud. De aquí que el nivel de cobertura de la población sea un determinante importante del nivel de utilización de los servicios de salud.²

El presente trabajo se fija básicamente dos objetivos. El primero es analizar cómo se determina la cobertura en salud, para lo que se explora la relación que existe entre ésta y distintas características socioeconómicas y demográficas de las personas. Con este enfoque se estudia tanto la decisión de contratar un seguro de salud como la de qué tipo de seguro contratar.

El segundo objetivo es presentar un método estadístico que permite predecir

* Agradezco la valiosa guía y colaboración de mi director, Walter Sosa Escudero, así como los comentarios y sugerencias de Omar Arias, Germán Coloma, Leonardo Gasparini, Alberto Porto y de los participantes del seminario del Departamento de Economía de la Universidad Nacional de La Plata. Por supuesto, los errores que pudieran subsistir son de mi exclusiva responsabilidad.

² Esto no es sólo una especulación teórica. La evidencia empírica en Argentina muestra claramente que el nivel de cobertura de la población es un determinante significativo de la demanda de servicios de salud. Ver *Bertranou (1998)* y *Convenio (1998)*.

cobertura utilizando información limitada. Si se contara con las encuestas apropiadas, la medición del nivel de cobertura de la población sería un ejercicio por demás simple. Sin embargo, debido a los elevados costos que suelen involucrar las encuestas, tanto las temáticas que abarcan como la frecuencia con que se realizan son usualmente reducidas. En particular, esto limita la disponibilidad de información actualizada sobre cobertura. El aporte metodológico propuesto permite predecir la cobertura de la población a partir de información usualmente disponible.

El ejercicio empírico consiste en la estimación de un modelo de la probabilidad de tener cobertura en base a información de la Encuesta Permanente de Hogares de 1992. Luego se usa este modelo para predecir la proporción de población con cobertura en el Área Metropolitana de Buenos Aires para el año 1996. A partir de estas estimaciones se analiza cómo la probabilidad de estar cubierto por un seguro de salud es afectada por cada uno de los determinantes que se identifican.

Condiciona la contratación de un seguro, también se estudia la determinación del tipo de cobertura. A tal fin se distinguen cuatro posibles formas de afiliación, usándose un modelo logístico multinomial para predecir las probabilidades de cada una de ellas y evaluar de qué manera los atributos de los individuos modifican las chances de tener un tipo u otro de cobertura.

El trabajo se organiza en tres partes. La Parte I se concentra en el análisis de la situación de cobertura (tenerla o no), mientras que en la Parte II se estudia la determinación de la forma de cobertura (tipo de seguro). Dentro de la Parte I se incluye una sección con el modelo teórico y la metodología empírica utilizados (Sección 2), otra que describe la información (Sección 3), otra que presenta las estimaciones (Sección 4), y una última en que se predice la población con cobertura para el Área Metropolitana de Buenos Aires en 1996 (Sección 5). La Parte II, a su vez, contiene una sección que describe el método empírico en base al cual se estudia la determinación de la forma de cobertura (Sección 6) y otra en que se presentan las estimaciones realizadas (Sección 7). El trabajo culmina con la Parte III destinada a las conclusiones y con un apéndice sobre los errores de predicción del modelo de la probabilidad de cobertura.

PARTE I: LA SITUACIÓN DE COBERTURA

En la Argentina el seguro de salud se brinda a través de entidades de distinto tipo (obras sociales, empresas de medicina pre-paga, mutuales, etc.). Las obras sociales son entidades que brindan seguro de salud a los trabajadores en relación de dependencia y a sus familiares directos. Las hay públicas y privadas. Una característica distintiva de las obras sociales es la obligatoriedad de la adhesión. El financiamiento de las obras sociales se basa en aportes de los trabajadores activos y pasivos (afiliados obligatorios), cuotas de afiliados voluntarios, co-seguros, etc.³

Es posible agrupar a las obras sociales según su dependencia administrativa en las que pertenecen al ANSSAL (Administración Nacional del Seguro de Salud), al INSSJP (Instituto Nacional de Servicios Sociales para Jubilados y Pensionados), otras obras sociales nacionales y obras sociales provinciales.

Entre las obras sociales nacionales comprendidas en el ANSSAL hay sindicales, estatales, por convenio, de personal de dirección, de administración mixta, etc. El Plan de Atención Médico Integral del INSSJP (PAMI) brinda cobertura en salud a los jubilados, pensionados y sus familiares. Hay otras obras sociales nacionales no nucleadas por el ANSSAL. Estas dan cobertura al personal del Poder Judicial, de las FF.AA y de Seguridad, del Congreso de la Nación y de las Universidades Nacionales. Por último, las obras sociales provinciales cubren a los trabajadores de la administración pública provincial y sus organismos descentralizados, y a sus familiares. En ciertos casos, también cubren a personal municipal.

El mercado privado de seguros de salud está conformado por las obras sociales privadas (pertenecientes al ANSSAL) y por empresas de seguros de contratación voluntaria. Entre los seguros voluntarios, los hay con y sin fines de lucro. Los primeros son los generalmente conocidos como planes de medicina pre-paga, mientras que los segundos suelen ser ofrecidos por mutualidades que cuentan con hospitales propios (Hospital Español, Italiano, etc.).⁴

2. El modelo

2.1. La decisión de cobertura

La naturaleza del seguro de salud en Argentina hace que puedan darse múltiples situaciones en relación a la cobertura de un individuo: tener seguro obligatorio, voluntario, ambos o ninguno, estar a cargo de un familiar afiliado o ser titular de la cobertura, estar afiliado a

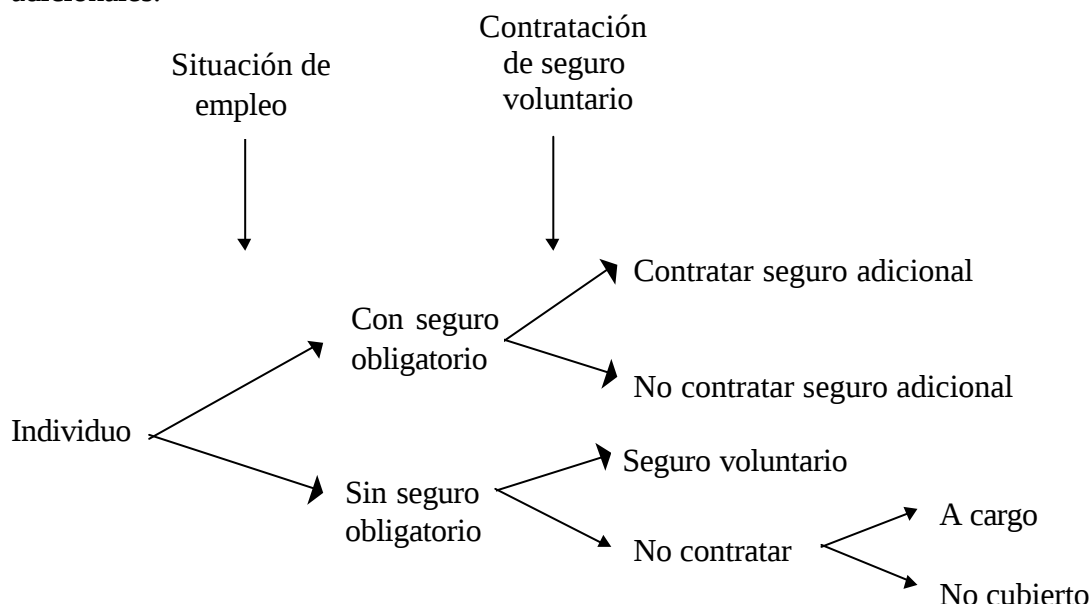
³ Sobre el financiamiento de las obras sociales, ver *Flood (1998)*

⁴ Para más información sobre el mercado privado de seguros de salud, ver *Flood (1998)* y *Bertranou (1998)*.

una obra social, a un sistema pre-pago o a otro tipo de plan, etc. La situación de empleo y la decisión de contratar voluntariamente un seguro son los elementos que esencialmente determinan cuál de todas estas posibilidades terminará caracterizando a un individuo en particular. Como fue señalado, las obras sociales cubren a los trabajadores en relación de dependencia y a sus familias en carácter obligatorio. De esta forma es evidente cómo el tipo de empleo determina la cobertura de una persona. El hecho de ser jubilado o pensionado también da indicios de la cobertura de una persona.

La contratación voluntaria, que se efectúa libremente en el mercado de seguros, puede hacerse para recibir cobertura en los casos en que no se tiene, o bien para complementar a otro seguro. Los trabajadores por cuenta propia, por ejemplo, no tienen un empleador que los provea obligatoriamente de un seguro de salud. Así, para obtener cobertura deben contratar voluntariamente uno. Por otra parte, ciertos individuos ya asegurados pueden estar deseosos de contratar otro seguro y de esta forma ampliar la cobertura ofrecida por el primero.

El diagrama a continuación hace algo más clara la discusión anterior.⁵ La situación de empleo conjugada con decisiones de contratación voluntaria puede conducir a una persona a alguna de las posibles formas de cobertura. Podría decirse que hay cuatro situaciones de cobertura posibles: asegurado obligatoriamente (titular de un seguro obligatorio), asegurado voluntariamente (titular de un seguro voluntario), a cargo de un familiar o sin cobertura. En los tres primeros casos, siempre puede haber seguros adicionales.



⁵ Con este gráfico no se quiere sugerir que la contratación de un seguro se hace a través de un proceso secuencial ni que la elección de empleo se hace en pos de la cobertura que el mismo brinda.

2.2. La decisión de cobertura según la teoría del consumidor

Según la teoría del consumidor se considera a los individuos como agentes maximizadores de la utilidad. En este contexto se espera que la elección hecha por un individuo sea aquella que le brinda el mayor nivel de utilidad entre todas las alternativas disponibles. Supongamos que un consumidor i enfrenta J alternativas. La utilidad de la elección de la j -ésima es U_{ij} . Por lo tanto se espera que si $U_{ij} > U_{ik}$ para todo $k \neq j$, el consumidor i elegirá la alternativa j . En el contexto de los seguros de salud que se está analizando, esto implica que la situación de cobertura que se observa para un individuo debería ser justamente la que maximiza su utilidad entre las alternativas posibles.

Sin embargo, y como más adelante se verá, hay un alto porcentaje de la población que hace un reducido o nulo ejercicio de elección. Ciertos miembros dentro de una familia “arrastran” a otros en sus decisiones y, por otra parte, una gran proporción de la población queda “atrapada” por el carácter de obligatoriedad de las obras sociales. Esto limita la posibilidad de interpretar la cobertura observada como el resultado de la elección de un individuo. La interpretación que cabría en este caso sería en términos de las características individuales y del hogar que hacen que un individuo sea más o menos “propenso” a tener un tipo de cobertura (incluyendo ninguna).

Más arriba se distinguieron cuatro posibles formas de cobertura: titular obligatorio, titular voluntario, a cargo de un familiar y sin cobertura. En esta parte del trabajo se analizarán los determinantes de tener o no cobertura, dejándose para la Parte II el estudio de la determinación de las distintas formas de afiliación. Así, a partir de ahora se tienen en cuenta sólo dos alternativas: cubierto y no cubierto.⁶

2.3. El método empírico

Una gran variedad de atributos personales hace que la situación de cobertura difiera entre individuos. Esos atributos pueden ser el nivel de ingresos, la aversión al riesgo, el sexo, el estado civil, la edad, el tamaño de la familia, el nivel educativo, el tipo de empleo, la propensión a enfermarse, etc. El conjunto de estas características determina que un individuo en particular sea más o menos propenso a tener cobertura. Sin embargo, el analista no observa todas las características susceptibles de influir sobre la propensión a tener cobertura en salud, ni el valor de la propensión en sí. La información disponible para el econométrico se limita a un subgrupo de las características relevantes y a un indicador de cobertura, que es el reflejo de la propensión no observada.

⁶ El análisis de los seguros complementarios no será encarado en este trabajo.

Pensemos en que la regresión subyacente es:

$$y_i^* = X_i \mathbf{b} + m_i$$

donde y^* es la propensión a tener cobertura en salud. Esta propensión viene determinada por los atributos incluidos en X_i , que es el vector de características observables del individuo i , de tamaño $(1 \times S)$. $\mathbf{b}$ es el vector de coeficientes de dimensión $(S \times 1)$ y m son perturbaciones aleatorias (todo aquello que afecta a la propensión a tener cobertura y que no es observado por el econometrista).

La propensión a tener un seguro de salud se refleja en el valor observado del indicador *cobertura* de la siguiente forma:

$$\begin{aligned} \text{cobertura}_i &= 1 \quad \text{si } y_i^* > 0 \\ \text{cobertura}_i &= 0 \quad \text{si } y_i^* \leq 0 \end{aligned}$$

Así, la probabilidad de que un individuo esté cubierto viene dada por:

$$\begin{aligned} \text{Prob}[\text{cobertura}_i = 1 \mid X_i] &= \text{Prob}[y_i^* > 0] \\ &= \text{Prob}[X_i \mathbf{b} + m_i > 0] \\ &= \text{Prob}[m_i > -X_i \mathbf{b}] \\ &= \text{Prob}[m_i < X_i \mathbf{b}]^7 \\ &= F[X_i \mathbf{b}] \end{aligned}$$

donde F es la función de distribución acumulada de m .

Esta formulación es común a los Modelos de Índices Lineales (Linear Index Models). La denominación es consecuencia de que la probabilidad del resultado observado es función del índice lineal $X_i \mathbf{b}$. La estimación de la probabilidad de cobertura implica encontrar estimaciones para $\mathbf{b}$. Cada elemento de este vector mide el efecto de un cambio marginal en un determinado atributo sobre el índice lineal, es decir, sobre la propensión a tener cobertura en relación a la posibilidad de no tenerla. A tal fin es necesario especificar cuál es la distribución de m . Si se supone que tiene una distribución de valores extremos, m tiene una distribución logística. Este modelo se conoce como Modelo Logit y será el que se use en el presente trabajo.⁸

⁷ Esto si la distribución de m_i es simétrica.

⁸ Otra alternativa hubiera sido suponer que m se distribuye normalmente. En este caso el modelo se conoce como Probit. Para un desarrollo de estos temas, ver *Greene (1997)*.

3. Los datos

3.1. Fuentes de información

La principal fuente de información con la que se cuenta es la Encuesta Permanente de Hogares (EPH) de mayo de 1992, para el Área Metropolitana, que incluye a la Capital Federal y al Gran Buenos Aires. Esta encuesta consta de dos partes: el cuestionario familiar y el individual. El primero está destinado a preguntas acerca de los hogares (características de la vivienda y del conjunto de personas que la habitan). El segundo se concentra en preguntas dirigidas a cada uno de los integrantes del hogar (características de las personas). Además se cuenta con el Anexo 1 de esta encuesta que tiene información individual relativa a los ingresos y beneficios sociales de las personas.

En 1992 se agregó al formulario usual de la EPH el “*Estudio sobre educación y utilización de servicios sociales*”. En este módulo se hacen preguntas relativas a la educación y asistencia alimentaria escolar, por un lado, y a los servicios de salud, por el otro. En particular, se efectuaron cuatro preguntas relativas a la afiliación a sistemas de atención de la salud. Las preguntas son las siguientes:

- 1) ¿Está Ud. afiliado a algún sistema de atención de la salud como obra social, mutual, pre-pago, etc.?.....SI o NO
- 2) ¿A cuales?...Nombres
- 3) ¿Es.....una obra social?
.....una mutual?
.....un pre-pago?
.....un sistema de emergencia?
.....un coseguro?
.....Otros?
- 4) ¿Está afiliado en carácter de titular o a cargo de un familiar? ¿Obligatorio o voluntario?

Gracias a este módulo es posible saber si una persona tiene o no cobertura de algún sistema de atención de la salud, las características de ese sistema y si la afiliación es en carácter de titular o a cargo de un familiar, así como si es voluntaria u obligatoria.

También se cuenta con la EPH del año 1996 para el Área Metropolitana de Buenos Aires. Dado que el módulo de educación y servicios de salud no se repitió, esta encuesta no cuenta con información relativa a la cobertura de las personas.

3.2. Descripción de la Base de Datos

Un hogar tipo en las EPH se presenta constituido por un jefe de hogar, un cónyuge e hijos. Muchas veces aparecen otros familiares como abuelos, hijos políticos, etc. Observando la

situación de los integrantes de un hogar respecto de la afiliación a un sistema de atención de la salud, se encuentra un espectro bastante amplio de casos. Hay familias en las que todos tienen cobertura, otras en las que sólo una parte la tiene, y otras en las que ninguno tiene cobertura. En los dos primeros casos, se da que algún o algunos integrantes son titulares del servicio y otros están a su cargo. Para realizar la estimación es importante determinar primero quiénes son los que realmente deciden si afiliarse o no en un hogar. Como se verá más adelante, y dando un ejemplo en extremo simplificado, es típico que el jefe de familia sea el afiliado titular y que automáticamente sus hijos menores de 18 años pasen a ser afiliados a su cargo. En este caso las características personales relevantes en la determinación de la situación de cobertura serían las del jefe de familia y no las de sus hijos menores, si bien la existencia o no de tales hijos podría ser un determinante importante de tal situación. Algo similar a lo que ocurre con los hijos se da con el cónyuge. Por supuesto que al agregar otros miembros de la familia este tipo de análisis se complica.

En lo que sigue se describe la información disponible y se presentan algunos cuadros que ilustran la situación antes descrita. Como se verá, hay una fuerte evidencia que lleva a pensar que la decisión del jefe de familia (ya sea al elegir un empleo que brinda seguro obligatorio o al contratar voluntariamente uno) determina la condición de afiliación de sus hijos menores. Algo similar ocurre con respecto a los cónyuges. Por su parte, los jubilados y pensionados presentan un alto índice de cobertura independientemente del rol que tengan en el hogar. Para los demás componentes del hogar (hijos mayores de 18 años y otros no jubilados ni pensionados que no son cónyuges) no se verifica una relación tan clara.

La base de datos original de la EPH contiene 10838 observaciones (10838 individuos pertenecientes a 3076 hogares distintos). El 68.74% de los individuos está cubierto por algún tipo de sistema de atención de la salud (un total de 7450 observaciones). Esta información se resume en el Cuadro 1.

Cuadro 1

Situación de cobertura.

Área Metropolitana, 1992.⁹

	observaciones	%
con cobertura	7450	68.74
sin cobertura	3327	30.7
ns/nc	61	0.56
Total	10838	100

- *Cobertura y relación de parentesco*

⁹ ns/nc: no sabe o no contesta la pregunta.

En la encuesta se distinguen ocho relaciones de parentesco respecto del jefe de familia, que a los propósitos de este trabajo fueron reagrupadas en cuatro: Jefes de Familia, Cónyuges, Hijos e Hijos Políticos y Otros (que incluye hermanos, otros familiares, servicio doméstico y otros componentes). Esta partición de la muestra resulta muy útil a fines descriptivos ya que permite distinguir patrones de comportamiento diferenciales para cada grupo. El grupo más numeroso es el de los Hijos e Hijos Políticos que constituye el 43% de la muestra (4653 observaciones). Luego sigue Jefes de Hogar con el 28%, Cónyuges con el 21% y por último Otros.

En relación a la proporción de gente con cobertura, el mayor índice está entre los Jefes de Hogar y los Cónyuges, con un 76,4% y 75.6% cubierto, respectivamente. Le siguen los Hijos e Hijos Políticos con 62.6% y en último lugar los Otros con un 56.6%.

En cuanto a su carácter, la afiliación a un sistema de atención de la salud puede ser como titular del seguro o a cargo de un familiar. De las 7450 personas cubiertas, un 45% lo es en carácter de titular. Como resulta evidente del Cuadro 3, una amplia mayoría de los Jefes con cobertura, es titular del seguro. Esto contrasta con los altos porcentajes de afiliación a cargo de un familiar que se registran para Cónyuges e Hijos e Hijos Políticos (72% y 81%, respectivamente). Para Otros no se observa ningún patrón evidente.

Cuadro 2

*Situación de cobertura según relación de parentesco
Área Metropolitana, 1992.*

	con cobertura		sin cobertura		ns/nc	
	obs.	%	obs.	%	obs.	%
<i>Jefes de Hogar</i>	2351	76.43	707	22.98	18	0.59
<i>Cónyuges</i>	1685	75.63	539	24.19	4	0.18
<i>Hijos e Hijos Políticos</i>	2915	62.65	1714	36.84	24	0.52
<i>Otros Componentes</i>	499	56.64	367	41.66	15	1.7

Cuadro 3

*Carácter de afiliación según relación de parentesco.
Área Metropolitana, 1992.*

	titular		a cargo		ns/nc	
	obs.	%	obs.	%	obs.	%
<i>Jefes de Hogar</i>	2128	90.51	219	9.32	4	0.17
<i>Cónyuges</i>	462	27.42	1219	72.34	4	0.24
<i>Hijos e Hijos Políticos</i>	513	17.6	2386	81.85	16	0.55
<i>Otros Componentes</i>	263	52.71	233	46.69	3	0.6

Cuadro 4

Cobertura de los Cónyuges de acuerdo a la cobertura del Jefe de Hogar.

Área Metropolitana, 1992.

En número de observaciones

	Jefe de Hogar cubierto		Total
	SI	NO	
con cobertura	1627	58	1685
sin cobertura	61	478	539
ns/nc	1	3	4
Total	1689	539	2228

En porcentajes por situación de cobertura del Jefe

	Jefe de Hogar cubierto	
	SI	NO
con cobertura	96.33%	10.76%
sin cobertura	3.61%	88.68%
ns/nc	0.06%	0.56%
Total	100.00%	100.00%

En el Cuadro 4 se presenta información acerca de la cobertura de los Cónyuges discriminada según el correspondiente Jefe de Hogar esté o no asegurado. Explorando la relación que existe entre la cobertura de estos dos grupos se observa que la probabilidad de que un Cónyuge esté cubierto condicional a que el Jefe de Hogar lo esté es del 96%, mientras que la probabilidad de que no esté cubierto si el jefe no lo está es casi del 89%.

- *Cobertura y edad*

Al igual que la partición según parentesco, el análisis de la cobertura por edades resulta muy ilustrativo.

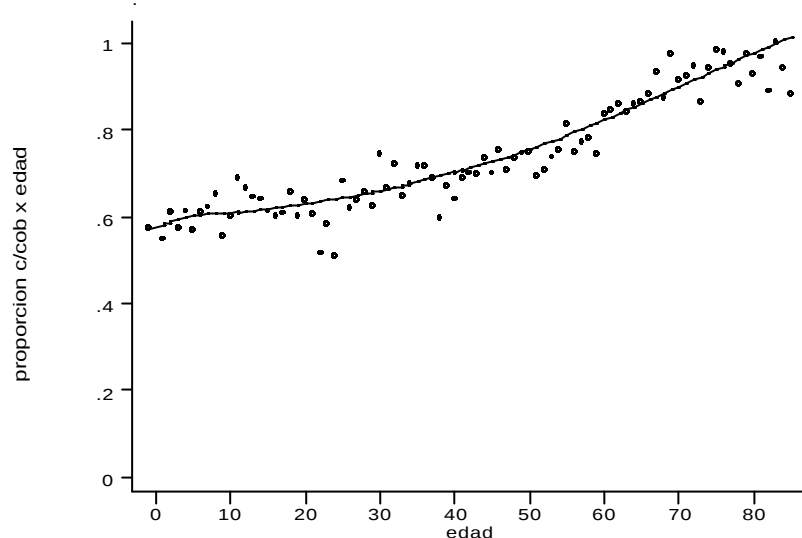
La nube de puntos del Gráfico 1 representa la proporción de gente con cobertura para cada edad y la línea continua es la regresión no paramétrica correspondiente.¹⁰ Como puede apreciarse claramente, existe una marcada relación positiva entre la edad y el nivel de cobertura en la muestra analizada.

Gráfico 1

Proporción de gente con cobertura de acuerdo a la edad

¹⁰ En relación a técnicas no paramétricas, ver *Deaton (1997)*.

Área Metropolitana, 1992



En el Cuadro 5 se agruparon los individuos en seis categorías en base a la edad. La proporción de gente cubierta en cada grupo aumenta con la edad, como sugería el gráfico anterior (desde algo más del 60% para los menores de 18 años hasta un 92% para los mayores de 65 años).

La alta proporción de cubiertos entre los mayores de 65 años se debe principalmente a que más del 80% de estas personas son jubilados o pensionados, grupo que tiene un índice de cobertura del orden del 98%.

Cuadro 5

Proporción de gente cubierta según edades
Área Metropolitana, 1992.

Edad	con cobertura	%	Total
Menores de 18	2291	61.27	3739
de 18 a 30	1163	61.50	1891
de 30 a 40	945	67.31	1404
de 40 a 50	958	71.81	1334
de 50 a 65	1050	78.42	1339
Mayores de 65	1043	92.22	1131

Hay 3739 menores de 18 años, de los que el 61% tiene cobertura en salud. El 97% de los jóvenes menores de 18 años con cobertura está a cargo de un familiar. A partir de esto es posible inferir que en casi todos los casos en que un menor de 18 años tiene cobertura hay alguien en su familia que también la tiene.

Investigando algo más esta evidencia, se encuentra una fuerte vinculación entre la cobertura de un menor de 18 años y la del Jefe del Hogar. Del Cuadro 6 se desprenden dos resultados interesantes: por un lado, de los jóvenes menores de 18 años cuyo Jefe de Hogar tiene cobertura, un 91% está cubierto. Es decir, la probabilidad de que un joven de estas características esté cubierto, condicional a que el Jefe de su Hogar lo está, es del 91%. Por otra parte, la probabilidad de que un joven menor de 18 años no esté cubierto, condicional a que el Jefe de su Hogar tampoco lo está, es del 92%.

Cuadro 6

*Cobertura de los menores de 18 años de acuerdo a la cobertura del Jefe de Hogar
Área Metropolitana, 1992.*

En número de observaciones			
	Jefe de Hogar cubierto		
	SI	NO	Total
con cobertura	2202	89	2291
sin cobertura	213	1219	1432
ns/nc	6	10	16
Total	2421	1318	3739

En porcentajes por situación de cobertura del Jefe

	Jefe de Hogar cubierto	
	SI	NO
con cobertura	90.95%	6.75%
sin cobertura	8.80%	92.49%
ns/nc	0.25%	0.76%
Total	100.00%	100.00%

- **Cobertura y sexo**

La composición por sexos en la muestra es aproximadamente igualitaria. El porcentaje de cobertura es también aproximadamente el mismo para hombres y mujeres: 68% y 69%, respectivamente. Esto indica que el sexo no sería un elemento relevante en la explicación de la probabilidad de estar cubierto. Sin embargo, el sexo sí parece ser determinante del carácter de la afiliación: mientras que un 60% de los hombres cubiertos son titulares del seguro, sólo el 32% de las mujeres cubiertas lo son.

Este análisis se enriquece cuando se agrega una tercera dimensión: la relación de parentesco. Como puede verse en el Cuadro 7, los Jefes de Hogar, ya sean hombres o mujeres, son en su gran mayoría titulares (92% de los Jefes hombres cubiertos y 86% de las Jefas mujeres cubiertas). Los Cónyuges hombres también presentan un alto porcentaje de titulares (88%), mientras que sólo el 26% de las Cónyuges mujeres cubiertas son titulares.

Es decir, ser hombre y jefe de hogar aumenta las chances de ser titular del seguro de salud.

Cuadro 7

Carácter de afiliación según parentesco y sexo.

Área Metropolitana, 1992.

	titular		a cargo	
	obs.	%	obs.	%
<i>Jefes Hombres</i>	1669	91.8	147	8.09
<i>Jefes Mujeres</i>	459	86.12	72	13.51
<i>Cónyuges Hombres</i>	23	88.46	3	11.54
<i>Cónyuges Mujeres</i>	439	26.46	1216	73.3

- *Cobertura y fuente de ingresos*

Se distinguen cinco fuentes de ingresos a partir de una reagrupación de las categorías presentadas en la Encuesta Individual de la EPH. Ingresos provenientes del trabajo asalariado (incluye categorías 1, 7, 8, 9 y 11), Ingresos provenientes de jubilaciones y/o pensiones (incluye categorías 5, 10 y 14), Ingresos de trabajadores por cuenta propia (categorías 2, 12, 13 y 15), Otras fuentes de ingresos, tales como rentas, alquileres, utilidades, etc. (incluye todas las demás categorías excepto la 23) y Sin ingresos (categoría 23).

En el Cuadro 8 se presenta para cada grupo de fuente de ingresos la cantidad de personas cubiertas y la proporción que las mismas representan dentro del mismo. También se lista la cantidad de personas con seguro obligatorio y aquellas afiliadas como titulares, y se expresan estas cifras en porcentajes respecto del total de cubiertos de cada categoría.

Para los tres primeros grupos se observa un claro patrón de comportamiento en relación a la cobertura. Los asalariados presentan un índice de cobertura del 76%, que resulta bastante superior al de la media de la muestra (68.7%). La gran mayoría tienen un seguro obligatorio y son titulares del mismo. Algo similar pero más acentuado sucede para aquellos con ingresos provenientes de jubilaciones y pensiones. En contraste, los trabajadores por cuenta propia tienen un bajo índice de cobertura (44%). Además, aproximadamente la mitad de los asegurados lo son a cargo de un familiar o como afiliados voluntarios.

Un alto porcentaje de cubiertos se registra también para los perceptores de ingresos de Otras fuentes (77%) con una proporción también alta de afiliados voluntarios en carácter de titular.

El rasgo más saliente y obvio en relación a los que no perciben ingresos es el bajísimo índice de titularidad (menos de un 3% de las personas cubiertas).

Cuadro 8

*Cobertura según fuente de ingresos.
Área Metropolitana, 1992.*

	con cobertura		titulares		obligatorios	
	obs.	%	obs.	%	obs.	%
<i>Asalariados</i>	2239	75.85	1841	82.22	1968	87.90
<i>Jubilados y Pensionados</i>	1159	98.14	1049	90.51	1101	95.00
<i>Cuentapropistas</i>	402	44.13	195	48.51	199	49.50
<i>Otras Fuentes</i>	278	77.22	182	65.47	135	48.56
<i>No perciben Ingresos</i>	3372	62.05	99	2.94	2526	74.91

4. Estimación de la probabilidad de cobertura

Haciendo uso de la EPH del Área Metropolitana para mayo de 1992 se estima un modelo logit como el descrito en la sección 2, prediciendo la probabilidad de estar cubierto para cada individuo de la muestra. El segundo paso es predecir las probabilidades de cobertura para 1996, año para el cual no se cuenta con el módulo de educación y servicios sociales. Esto se hace utilizando los parámetros estimados para 1992 y la información de la EPH de octubre de 1996 para el Área Metropolitana. El supuesto detrás de esta metodología es que los parámetros del índice lineal son constantes en el tiempo. Cabe esperar que ante la ausencia de grandes reformas estructurales en el área de los seguros de salud, la forma en que se determina la cobertura de la gente no presente variaciones significativas en el tiempo, por lo que este supuesto no parece ser demasiado fuerte.¹¹

Una vez predicha la probabilidad de cobertura para los individuos de los distintos aglomerados, y haciendo uso de los factores de ponderación correspondientes, se traducen las probabilidades a niveles de cobertura de la población.

La estimación se lleva a cabo en esta sección mientras que las predicciones para el año 1996 se dejan para la sección 5.

4.1. Restricción de la muestra

Una restricción a la muestra original se impone al incluir al ingreso como variable

¹¹ Ha habido cierta reorganización de las obras sociales que no debería afectar sustancialmente el lado de la demanda. Lo que sí ha experimentado una expansión significativa en este período es el mercado privado de seguros de salud. La desregulación de este sector llevó a un desarrollo importante de sistemas pre-pagos y de emergencias. Seguramente el aumento de la competencia haya llevado a una reducción de los precios y/o a un aumento en la calidad del servicio brindado en este sector. Si bien estas modificaciones deberían tener el efecto de fomentar a la demanda, parecería que las personas susceptibles de contratar un seguro de este tipo ya habrían estado aseguradas desde antes. Con esta consideración, si bien es factible que los patrones de afiliación se hayan modificado, no parecería que esto debería afectar significativamente al nivel de cobertura de la población.

explicativa (ver subsección 4.2.). Debido a inconsistencias en el reporte del ingreso individual (ocupados o jubilados que dicen no percibir ingresos, etc.) no siempre puede confiarse en esta medida. Por esta causa, se eliminan de la muestra todas las observaciones para las que el INDEC ha detectado tales inconsistencias.

Como ya fue señalado, para llevar adelante la estimación del modelo es necesario identificar a aquellos individuos cuyas características personales determinan la situación de cobertura dentro de cada hogar. Del análisis anterior se desprende que tanto la cobertura de los jóvenes menores de 18 años como la de los Cónyuges parece depender de la de los Jefes de Hogar y que puede ser inferida a partir de ella. Es posible hacer esta inferencia usando el Teorema de Bayes, que permite conocer la probabilidad de un evento a partir de su probabilidad condicionada en la ocurrencia de otro.

La formulación del Teorema de Bayes es la siguiente:

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)}$$

de donde,

$$P(B) = \frac{P(B|A) \cdot P(A)}{P(A|B)}$$

En términos del modelo de cobertura esta formulación implica que, por ejemplo, la probabilidad de un Cónyuge de estar cubierto por un seguro de salud es igual a la probabilidad que esté cubierto condicional a que el Jefe de su Hogar lo está multiplicada por la probabilidad de que el Jefe del Hogar esté cubierto, dividida por la probabilidad de que el Jefe de Hogar esté cubierto condicional en que el Cónyuge lo está. La probabilidad de que el Jefe del Hogar esté cubierto se estima usando un modelo Logit. Las dos probabilidades condicionales se calculan a partir del Cuadro 4. Esto último constituye una simplificación que surge de suponer constantes estas relaciones para todas las familias.

Una vez que se eliminan las observaciones con ingresos incoherentes y estos dos grupos de la muestra original, todos los jubilados y pensionados tienen cobertura, es decir, la característica Jubilado o Pensionado predice perfectamente la situación de cobertura. Por esto, se eliminan todas las observaciones correspondientes a Jubilados y Pensionados de la muestra. Como se dijo antes, la probabilidad *ex ante* de estar cubiertos es del 98.2% para este grupo, independientemente de la cobertura de cualquier otro miembro de su hogar. Esta es la probabilidad de estar cubierto que en este trabajo se le asigna a un jubilado o pensionado.

En resumen, se eliminan cuatro grupos por tres motivos distintos. Los que tienen ingresos incoherentes, los Cónyuges, los menores de 18 años, y los Jubilados y

Pensionados. La causa de la eliminación del primer grupo es la falta de confianza que se puede tener de la información disponible. En cuanto a los Cónyuges y menores de 18 años, se considera que su situación de cobertura depende de una elección hecha por el Jefe de Hogar correspondiente. Por último, ser Jubilado o Pensionado es una característica que por sí sola determina la tenencia de un seguro de salud, y por lo tanto no es posible incluirla en el modelo de la probabilidad de tener cobertura.

La muestra reducida tiene 2414 observaciones. Un 69.43% de las personas tiene cobertura, con lo que la proporción de asegurados es prácticamente la misma que la de la muestra original. De los afiliados a un sistema de atención de la salud, un 81% son titulares. Este porcentaje difiere significativamente del 45% de la base original. La causa de esta gran diferencia está en la eliminación de los jóvenes menores de 18 años y de los Cónyuges, que son grupos con una amplia mayoría de cubiertos a cargo de un familiar. También la composición por sexos se modifica significativamente. Mientras que en la muestra original la participación de mujeres y hombres era aproximadamente igualitaria, en la muestra reducida un 72% son hombres. Esto se debe sobre todo a la eliminación de los Cónyuges (que son mayoritariamente mujeres).

En el Cuadro 9 se presentan las cifras que describen a la muestra reducida.

Cuadro 9

*Cobertura en la muestra reducida.
Área Metropolitana, 1992.*

<i>Situación de cobertura</i>		
	<i>observaciones</i>	<i>%</i>
con cobertura	1676	69.43
sin cobertura	738	30.57
total	2414	100.00

<i>Tipo de cobertura</i>		
	<i>observaciones</i>	<i>%</i>
titular obligatorio	1177	70.23
titular voluntario	180	10.74
a cargo	311	18.56
ns/nc	8	0.48
total	1676	100.00

4.2. Descripción de las variables del modelo

Teniendo en cuenta que uno de los objetivos es predecir cobertura cuando no se cuenta con información de afiliación a sistemas de atención de la salud, no pueden incluirse datos de

cobertura como variables explicativas del modelo a estimar. Por ejemplo, si se usara como variable explicativa la cobertura de un determinado miembro de la familia, no sería posible predecir probabilidades en aglomerados que no cuenten con información de cobertura. Por este motivo se opta por especificar un modelo en forma reducida.

La variable dependiente en este modelo toma dos posible valores: *cubierto=1* y *cubierto=0*, indicando que la persona está o no cubierta, respectivamente.

Las variables explicativas que se incluyen en el modelo a estimar son todas aquellas que se consideran relevantes para la decisión de cobertura y que aparecerían en la forma reducida del modelo. A continuación se las describe detalladamente.

Variables demográficas

- 1 *edad*
- 2 *sexo*: variable binaria que toma el valor uno para los hombres y cero para las mujeres.
- 3 *casado*: variable binaria que toma el valor uno para los casados y cero para los demás estados civiles.
- 4 *menores*: indica la cantidad de hijos e hijos políticos menores de 18 años que tiene una persona. Es indicativa del tamaño de la familia.
- 5 *indicadoras de la relación de parentesco*: son cuatro variables binarias que indican la relación de parentesco respecto del Jefe del Hogar: *Jefe=1* (Jefe del Hogar), *Cónyuge=1* (Cónyuge), *Hijo=1* (Hijos e Hijos Políticos) y *Otrarel=1* (Otros). Obviamente *Cónyuge* no se incluye dado que esta categoría se elimina al formar la muestra reducida.

Variables indicativas de la condición laboral

- 6 *empleo*: variable binaria que toma el valor uno para todas las personas ocupadas y cero en los demás casos.

Variables relacionadas a los ingresos

- 7 *liaee*: logaritmo natural del ingreso ajustado por adulto equivalente y economías de escala. Es una medida alternativa del ingreso per capita que considera de una forma distinta a cada miembro de la familia de acuerdo a su edad y sexo y ajusta el tamaño del hogar para reconocer economías de escala internas al mismo.¹²
- 8 *fuentes de ingresos*: son cinco categorías que indican la fuente de ingresos: *asal=1* (trabajo asalariado), *fjub=1* (jubilación o pensión), *cuentap=1* (trabajo por cuenta

¹² En *FIEL (1998)* puede verse cómo se construye esta variable.

propia), *otrasf*=1 (otras fuentes, tales como rentas, alquileres, utilidades, etc.) y *noperci*=1 (no perciben ingresos). Antes se aclaró que los jubilados no forman parte de la muestra reducida. Sin embargo, dado que las categorías aquí presentadas surgen de reagrupar otras, pueden darse casos en que si bien la persona percibe ingresos de jubilaciones o pensiones, no se reporta como inactiva porque también hace, por ejemplo, trabajos por cuenta propia. Así es que, pese a que los Jubilados y Pensionados inactivos son eliminados para formar la muestra reducida, quedan personas que perciben algunos ingresos de jubilaciones o pensiones.

Variables relacionadas a la educación

9 *Indicativas del nivel educativo*: son siete categorías que indican el máximo nivel de educación alcanzado por un individuo: *Preesc*=1 (preescolar), *Primi*=1 (primaria incompleta), *Primc*=1 (primaria completa), *Seci*=1 (secundaria incompleta), *Secc*=1 (secundaria completa), *Univi*=1 (superior o universitaria incompleta) y *Univc*=1 (superior o universitaria completa). *Primi* será la categoría base u omitida en el análisis.

En el Cuadro 10 se presentan los estadísticos descriptivos de cada una de estas variables discriminados según situación de cobertura y para el total de la muestra restringida, respectivamente.

Cuadro 10

Estadísticos descriptivos de las variables.

Muestra restringida. Área Metropolitana, 1992.

	SIN COBERTURA				CON COBERTURA				TOTAL			
Variable	Media	D. Std.	Min	Max	Media	D. Std.	Min	Max	Media	D. Std.	Min	Max
edad	36.56	13.74	19	88	38.91	13.85	19	90	38.19	13.86	19	90
sexo	0.67	0.47	0	1	0.74	0.44	0	1	0.72	0.45	0	1
casado	0.37	0.48	0	1	0.57	0.50	0	1	0.51	0.50	0	1
jefe	0.59	0.49	0	1	0.70	0.46	0	1	0.66	0.47	0	1
hijo	0.32	0.47	0	1	0.25	0.43	0	1	0.27	0.44	0	1
otrarel	0.09	0.29	0	1	0.05	0.22	0	1	0.06	0.25	0	1
menores	1.03	1.59	0	9	0.94	1.31	0	8	0.97	1.40	0	9
empleo	0.70	0.46	0	1	0.89	0.31	0	1	0.83	0.37	0	1
liaee	5.41	0.67	3.14	7.33	5.87	0.70	3.14	8.19	5.73	0.72	3.14	8.19
asal	0.41	0.49	0	1	0.73	0.44	0	1	0.63	0.48	0	1
fjub	0.00	0.00	0	0	0.04	0.20	0	1	0.03	0.17	0	1
otrasf	0.04	0.19	0	1	0.05	0.23	0	1	0.05	0.22	0	1
noperci	0.26	0.44	0	1	0.09	0.28	0	1	0.14	0.34	0	1
preesc	0.00	0.04	0	1	0.00	0.00	0	0	0.00	0.02	0	1
primc	0.42	0.49	0	1	0.29	0.45	0	1	0.33	0.47	0	1
seci	0.20	0.40	0	1	0.20	0.40	0	1	0.20	0.40	0	1
secc	0.14	0.35	0	1	0.19	0.39	0	1	0.17	0.38	0	1
univi	0.05	0.22	0	1	0.14	0.35	0	1	0.11	0.32	0	1
univc	0.02	0.14	0	1	0.10	0.30	0	1	0.08	0.27	0	1
obs.	738				1676				2414			
% del total	30.57				69.43				100			

4.3. Estimación del modelo

Usando la muestra reducida se estima el siguiente Modelo Logit de la probabilidad de cobertura:

$$Prob(cubierto=1)=F(X \mathbf{b})$$

donde $\mathbf{b}$ es el vector de coeficientes y X es el vector de variables explicativas.

Para estimar este modelo se usa el paquete econométrico Stata, versión V. La salida de la regresión se presenta en el Cuadro 11. En la primera columna aparecen las variables del modelo mientras que en la segunda se reportan los coeficientes estimados para cada una. Los desvíos estándar estimados de los coeficientes y corregidos por heterocedasticidad aparecen en la columna tres, y el cociente entre los coeficientes y estos últimos se presenta en la columna cuatro.¹³ La última columna reporta el valor “p” asociado al coeficiente. Ese valor de probabilidad permite determinar si el coeficiente es o no significativamente distinto de cero, es decir, si la variable correspondiente tiene realmente un efecto significativo sobre la probabilidad de cobertura. Para un nivel de significación del 5%, se puede considerar como significativamente distinto de cero a todos aquellos coeficientes para los cuales el valor de probabilidad reportado en la columna cinco sea menor a 0.05.

Si bien la muestra restringida cuenta con 2414 observaciones, sólo se utilizan 2343

¹³ Stata calcula errores estándar de Huber. Ver *Greene (1997)*.

para la regresión. Del Cuadro 10 surge que tanto *fjub* como *preesc* predicen perfectamente la situación de cobertura. De hecho, todos los individuos de la muestra reducida que reciben ingresos de jubilaciones o pensiones tienen cobertura, y ninguno de los que tienen como máximo nivel educativo el preescolar está cubierto. Este es un caso de predictores perfectos, por lo que se excluyen ambas variables y por consiguiente todas las observaciones para las que toman el valor uno. Así, se eliminan un total de 71 observaciones (la diferencia entre 2414 y 2343).

Según el estadístico Chi cuadrado reportado en la parte superior derecha de la tabla, las variables incluidas en este modelo son conjuntamente significativas para explicar la decisión de cobertura de los individuos.

En relación a los efectos estimados de las variables, prácticamente todos van en la dirección esperada. En los modelos lineales el coeficiente mide el efecto marginal de la variable independiente en la dependiente. Dado que el Logit no es un modelo lineal, los coeficientes del Cuadro 11 no pueden ser interpretados de esta manera. A pesar de este problema, aún es posible conocer la dirección del efecto y su magnitud relativa observando los signos y el tamaño relativo de los coeficientes.

La edad ejerce un efecto positivo sobre la probabilidad de tener un seguro de salud. Es necesario tener en cuenta que este resultado no es consecuencia de que los jubilados sean personas mayores con un muy alto índice de cobertura, ya que este grupo no se incluye en la muestra reducida. Probablemente este efecto esté reflejando un aumento, con la edad, de la necesidad de asistencia médica o de la aversión al riesgo de las personas.

El sexo no afecta significativamente la situación de cobertura una vez que se controla por el rol que cada individuo desempeña en el hogar y por su estado civil entre otras cosas. Respecto de esta última variable, se observa que los casados tienen una mayor probabilidad de tener un seguro de salud que los solteros, viudos, unidos, separados o divorciados.

En cuanto a las relaciones de parentesco, los cónyuges no entran en el análisis y la categoría omitida en relación a la cual corresponde interpretar los coeficientes es la de jefes de hogar. Del Cuadro 11 se concluye que ni ser *hijo* ni ser *otrarel* modifican la probabilidad de contratar un seguro en comparación a la de un jefe. Esto indicaría que las diferencias en las probabilidades de cobertura que existen entre individuos de la muestra restringida no son originadas en los distintos roles que cada uno desempeña en el hogar.

Cuadro 11

*Estimación del modelo Logit de la probabilidad de estar cubierto.
Área Metropolitana, 1992.*

Obs. =2343 chi2(16) =489.92 Prob > chi2 =0.0000 Pseudo R2 =0.2538 Log Likelihood = -1088.6335					
cubierto	Coef.	D. Std. *	z	P> z	Odds Ratio
edad	0.0221	0.0051	4.2860	0.0000	1.0223
sexo	-0.1790	0.1455	-1.2300	0.2190	0.8361
casado	1.1464	0.1399	8.1910	0.0000	3.1467
hijo	0.1694	0.1920	0.8830	0.3770	1.1846
otrarel	-0.2181	0.2270	-0.9600	0.3370	0.8041
menores	-0.0596	0.0491	-1.2140	0.2250	0.9421
empleo	0.5010	0.3252	1.5410	0.1230	1.6504
liaee	0.6578	0.0955	6.8850	0.0000	1.9305
asal	2.3750	0.1472	16.1380	0.0000	10.7513
otrasf	1.4383	0.3234	4.4470	0.0000	4.2134
noperci	1.2196	0.3582	3.4050	0.0010	3.3859
primc	0.4183	0.1771	2.3620	0.0180	1.5194
seci	1.0882	0.2082	5.2270	0.0000	2.9690
secc	1.1529	0.2127	5.4190	0.0000	3.1674
univi	2.3953	0.2835	8.4480	0.0000	10.9720
univc	2.1730	0.3607	6.0240	0.0000	8.7843
constante	-7.1384	0.6919	-10.3170	0.0000	

* Corregidos por heterocedasticidad

Nota: fjub=1 predice perfectamente cobertura. fjub no se considera y 70 obs. se eliminan
preesc=1 predice perfectamente no cobertura. preesc no se considera y 1 obs. se elimina.

Cuadro 12

Tests sobre los coeficientes del modelo Logit.

Tests de Wald

Ho: los coeficientes son iguales		
seci secc	chi2(1) = 0.14	Prob > chi2 = 0.7112
univi univc	chi2(1) = 0.35	Prob > chi2 = 0.5546
hijo otrarel	chi2(1) = 3.14	Prob > chi2 = 0.0763
asal otrasf	chi2(1) = 9.02	Prob > chi2 = 0.0027
otrasf noperci	chi2(1) = 0.38	Prob > chi2 = 0.5361

Se puede pensar que la influencia de *menores* sobre la probabilidad de tener cobertura se debe a que una mayor cantidad de hijos hace que el riesgo financiero de las enfermedades sea mayor, y por tanto mayor la probabilidad de contar con un seguro. Sin embargo, según la estimación y contrario a la intuición anterior, el número de Hijos e Hijos Políticos menores de 18 años no tiene un efecto significativo sobre la probabilidad de contratar un seguro de salud en la muestra restringida.

Por su parte, los aumentos en el ingreso ajustado por adulto equivalente y economías de escala contribuyen a un aumento de la probabilidad de estar cubierto.

Además, la fuente de donde provienen los ingresos puede determinar que aunque ese ingreso sea el mismo para dos individuos (y manteniendo constante todo lo demás), la probabilidad de estar cubierto difiera entre ambos. Cabe recordar que la categoría *ffub* no se incluye en el análisis. En este caso la categoría base contra la cual corresponde hacer la interpretación de los coeficientes es *cuentap*. La probabilidad de que un cuentapropista tenga seguro de salud es menor que la de cualquiera de las tres categorías incluidas en el análisis. Así, la probabilidad de estar cubierto es máxima para los asalariados, siguen los que no perciben ingresos y los que lo hacen de otras fuentes (según el test reportado en el Cuadro 12, estos dos últimos efectos son significativamente iguales) y por último están los cuentapropistas. Estos resultados se vinculan con el hecho de que para los asalariados el seguro es generalmente provisto por el empleador y la afiliación es obligatoria. En cuanto a los cuentapropistas, la afiliación es voluntaria, así como para los que perciben ingresos de fuentes tales como rentas y beneficios. La diferencia de probabilidades estimada para estos dos últimos grupos se debe a características propias de cada uno que no son captadas por las otras variables del modelo.

Un resultado curioso es que la variable *empleo* no es significativa en la determinación de la probabilidad de cobertura. Esto probablemente se deba a que las variables de fuentes de ingresos dan una descripción más exhaustiva del empleo, captando así toda la influencia que este pueda tener sobre la situación de cobertura.

En relación a la educación se observa un resultado interesante. Como se señaló antes, la categoría *preesc* no se incluye en el análisis y la categoría base es *primi*. Para todos los grupos se estiman coeficientes positivos que crecen con el nivel educativo. Esto implica que a medida que la educación aumenta también lo hace la probabilidad de tener un seguro de salud.¹⁴ La excepción parece darse entre *univi* y *univc*. Sin embargo, según los tests del Cuadro 12, los efectos de *univi* y *univc* sobre la probabilidad de cobertura son significativamente iguales. Lo mismo ocurre con los de *seci* y *secc*. De esta forma se hace evidente la influencia positiva de la educación sobre la probabilidad de tener cobertura en salud. Es probable que la educación actúe sobre la percepción del riesgo de las personas o sobre su aversión al mismo y que por ese motivo influya en la cobertura.

En resumen, diferencias inter-personales en edad, estado civil, nivel y fuente de ingresos y nivel educativo provocarían diferencias en las probabilidades de contratar un seguro de salud. Ninguna de las demás variables incluidas en el modelo contribuyen a explicar la situación de cobertura.

Hasta ahora pudo inferirse la dirección e importancia relativa de los efectos que las

¹⁴ En el extremo se tenía que *preesc*=1 predecía perfectamente no estar cubierto y por ese motivo no se estimó este modelo para ese grupo.

distintas variables tienen sobre la probabilidad de cobertura. Para determinar la magnitud de los mismos es necesario recurrir a ciertos cálculos auxiliares. Así, el efecto marginal de una variable continua sobre la probabilidad de cobertura viene dado por:

$$\frac{\partial F}{\partial x_k} = \beta_k \cdot f(X\beta)$$

donde x_k es una variable continua y $f(.)$ es la función de densidad logística.

De esta expresión resulta evidente que el efecto marginal depende de “dónde” se lo mida, es decir, del valor que se le asigne a las variables explicativas. En general el efecto marginal se evalúa en la media de las variables continuas y para valores fijos arbitrarios de las variables dicotómicas.

En el Cuadro 13 se reportan los efectos marginales de la edad y el ingreso sobre la probabilidad de contratar un seguro de salud. Las estimaciones anteriores arrojaron que el sexo no es significativo, motivo por el cual los efectos marginales para uno y otro grupo no difieren significativamente. Pese a esto, en el cuadro se presentan las cifras para hombres y mujeres. Debido a que las medias de las variables continuas difieren entre sexos, el efecto marginal se mide en “otro lugar” según se trate de hombres o de mujeres. Por esta causa y a fines ilustrativos se decidió presentar las cifras para los dos grupos. Para cada sexo se analizan seis subgrupos distintos: asalariados y cuentapropistas con tres posibles niveles de educación máxima alcanzada (primaria completa, secundaria completa y universitaria o superior completa), todos ellos jefes de hogar casados. Para los asalariados, independientemente del sexo, se observa que tanto el efecto de la edad como del ingreso sobre la probabilidad de contratar cobertura disminuye a medida que aumenta el nivel educativo. Un año más aumentaría en 0.3% la probabilidad de tener cobertura para un hombre con educación primaria y con edad, número de hijos menores e ingreso promedio. Por su parte, este aumento sería sólo de 0.07% para un universitario (un 24% del anterior). Un aumento de un 1% del ingreso ajustado por adulto equivalente y economías de escala produce un aumento del 9% en la probabilidad de tener cobertura de un hombre con educación primaria, y del 2% para uno con educación universitaria (lo que implica que el efecto se reduce en un 78% al pasar de *primc* a *univc*).¹⁵ Para los hombres con estudios secundarios se verifica un efecto de magnitud intermedia, dos veces y media superior al que se registra para universitarios, o casi un 60% que el de los primarios. Para las mujeres se observan relaciones muy similares.

Para los cuentapropistas este patrón se modifica. En primer lugar, si bien el efecto

¹⁵ Es posible analizar el cambio en la probabilidad para variaciones porcentuales en el ingreso debido a que la variable explicativa es el logaritmo del mismo.

marginal disminuye si se compara un individuo con educación primaria con un universitario, la diferencia es menor que la que se verificaba para los asalariados. Por ejemplo, un año más para los hombres universitarios tiene un 90% del efecto que para los que tienen como máximo una educación primaria. En segundo lugar, el efecto marginal de las dos variables analizadas es mayor para el nivel educativo secundario completo que para el primario.

Cuadro 13

*Efectos marginales de la edad y el ingreso sobre la probabilidad de cobertura.
Jefes de Hogar casados.*

Hombres	asal		cuentap	
	edad	liaee	edad	liaee
primc	0.0029	0.0874	0.0049	0.1458
secc	0.0017	0.0498	0.0055	0.1644
univc	0.0007	0.0200	0.0044	0.1261

Mujeres	asal		cuentap	
	edad	liaee	edad	liaee
primc	0.0027	0.0811	0.0051	0.1508
secc	0.0015	0.0455	0.0055	0.1636
univc	0.0006	0.0181	0.0040	0.1195

* evaluado en las medias de *edad*, *menores* y *liaee*. Para los hombres las medias son 38.6, 1.2 y 5.7, respectivamente. Para las mujeres, los valores correspondientes son 34.8, 0.3 y 5.7.

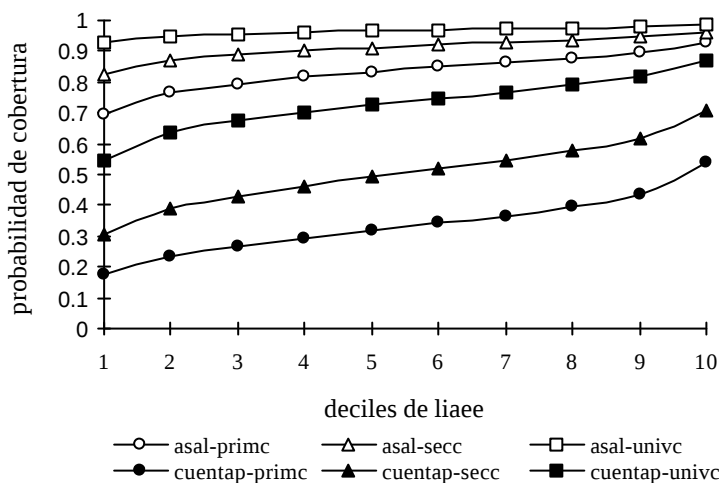
Haciendo uso del Gráfico 2 se puede analizar más detalladamente el efecto del ingreso sobre la probabilidad de contratar un seguro de salud de los hombres que son jefes de hogar y casados. Las tres curvas superiores son para asalariados y las tres inferiores para cuentapropistas. Cada una de las tres corresponde a un nivel educativo distinto: universitario, secundario y primario (de arriba hacia abajo). La altura de las mismas representa la probabilidad de que un individuo representativo de cada grupo tenga cobertura, mientras que su pendiente mide el efecto marginal de un cambio porcentual de su ingreso ajustado.

Lo primero que surge del gráfico es que las curvas nunca se cruzan. Esto es consecuencia de dos fenómenos. Primero, independientemente del nivel de ingresos y del máximo nivel educativo alcanzado, un asalariado siempre tiene mayor probabilidad de tener cobertura que un cuentapropista. Esto resulta coherente con el hecho de que la cobertura de los asalariados es generalmente obligatoria. Segundo, independientemente del nivel de ingresos y dada la fuente de procedencia de los mismos, a mayor nivel educativo, mayor probabilidad de contratar un seguro de salud. Como ya fue señalado, esto podría deberse a que la educación modifica la percepción que las personas tienen del riesgo.

Gráfico 2

Probabilidad de cobertura de acuerdo al nivel de ingresos.

Simulaciones para Jefes de Hogar sexo masculino y casados.¹⁶



Para los asalariados hay evidencia de un efecto marginal decreciente del ingreso sobre la cobertura para los tres niveles de educación, que se refleja en la concavidad de las curvas. Además, a medida que aumenta el ingreso se produce una especie de convergencia de la probabilidad de cobertura para los tres grupos de asalariados, como consecuencia de un más rápido crecimiento de la misma para niveles educativos inferiores. Esto indica que el efecto marginal del ingreso, si bien decreciente, es mayor cuanto menor es la educación. De esta forma podría interpretarse que aumentos sucesivos del ingreso tienden a reducir las diferencias en cobertura generadas por diferencias en el nivel educativo.

En cuanto a las curvas que representan a los cuentapropistas, las mismas no presentan evidencia alguna de convergencia hacia una misma probabilidad de cobertura. Al igual que para los asalariados, a mayor nivel educativo mayor es la probabilidad de contratar un seguro. Sin embargo, a diferencia de aquellos, para los cuentapropistas las brechas de probabilidades existentes para los distintos niveles educativos se mantienen ante variaciones del ingreso. Esto se debe a que, para este grupo de cuentapropistas, el efecto marginal del ingreso no varía demasiado con el nivel educativo.

De la comparación de asalariados y cuentapropistas puede decirse que si bien la probabilidad de cobertura es superior para los primeros, con los efectos marginales sucede

¹⁶ Las probabilidades se evalúan en las medias muestrales de *edad* y *menores*.

lo contrario. El ingreso tiene un mayor efecto en la decisión de cobertura de un cuentapropista que de un asalariado. Nuevamente, esto resulta obvio si se tiene en cuenta la naturaleza obligatoria del seguro de aquellos que están empleados en relación de dependencia.

Los efectos marginales para las variables binarias se calculan en forma diferente debido a que no es posible aplicar para ellas el concepto de derivada. Así, el efecto marginal sobre la probabilidad de cobertura cuando la variable binaria d_k pasa de valer cero a uno viene dado por:

$$EfMg d_k = F(b_k + Xb) - F(Xb)$$

donde F es la función de distribución acumulada logística y X tiene un cero en la posición k -ésima. Como se señaló antes, dado que el efecto marginal depende de los valores que toman cada una de las variables, se lo evalúa usualmente para los valores medios de las variables continuas y para valores fijados arbitrariamente de las dummies.

En el Cuadro 14 se reportan los efectos marginales de algunas de las variables cualitativas para distintos grupos de individuos. Si bien el sexo no resulta ser una variable significativa del modelo, se presentan los efectos sobre la probabilidad de contratar un seguro de esta variable, así también como de otras características desagregados por sexo. Esto se hace debido a que las medias de las variables continuas difieren entre hombres y mujeres, por lo que los resultados para un “representante promedio” podrían cambiar significativamente según se trate de uno u otro género.

De la primera parte del cuadro surgen algunos resultados interesantes. Para los cuatro grupos de individuos analizados, ser *hombre promedio* aumenta la probabilidad de tener cobertura en relación a una *mujer promedio*. Consistente con los hallazgos anteriores, los efectos marginales parecen acentuarse cuanto menor es el nivel educativo. En particular, aquí se aprecia que entre los Jefes de Hogar, casados y asalariados, el efecto del sexo sobre la probabilidad de cobertura es cuatro veces y media mayor para los que tienen un nivel primario que uno universitario. Dentro del grupo que tiene completo este último nivel educativo, el efecto del sexo es mayor para los solteros (casi tres veces más que para los casados) y cuentapropistas (seis veces y media mayor que para asalariados).

Cuadro 14

*Efectos marginales sobre la probabilidad de cobertura de variables binarias.**

Efectos del sexo (de mujer a hombre)

casado jefe asal primc	0.0227
casado jefe asal univc	0.0051
soltero jefe asal univc	0.0142
casado jefe cuentap univc	0.0331

Efectos del estado civil (de soltero a casado)

mujer jefe asal univc	0.0531
hombre jefe asal univc	0.0622

Efectos de la fuente de ingresos (de cuentapropista a asalariado)

mujer casada jefe univc	0.2021
hombre casado jefe univc	0.2302

* evaluado en las medias de *edad*, *menores* y *liaee*. Para los hombres las medias son 38.6, 1.2 y 5.7, respectivamente. Para las mujeres, los valores correspondientes son 34.8, 0.3 y 5.7.

La segunda parte del cuadro presenta los efectos sobre la probabilidad de cobertura de cambios en el estado civil. Como ya fue señalado, hay evidencia de que el estado civil afecta significativamente la decisión de cobertura de las personas. Para una “mujer promedio”, casarse aumenta su probabilidad de tener cobertura en 0.05. A fines ilustrativos se presenta también esta cifra para un “hombre promedio”. Cabe aclarar nuevamente que el efecto del sexo no es significativo.

En cuanto a magnitud, el efecto marginal de cambios en la fuente de ingresos es el de mayor importancia. Para Jefes de Hogar, casados y universitarios, pasar de ser cuentapropista a asalariado aumenta en 20% la probabilidad de contratar un seguro de salud. Esta magnitud es la que se obtendría de calcular la distancia vertical entre las dos curvas del Gráfico 2 que corresponden a educación universitaria, en el nivel medio de *liaee* (5.7 aproximadamente).

4.4. Predicción de la cobertura para la muestra reducida.

Usualmente se considera que aquellos individuos para los cuales la probabilidad de cobertura estimada supera el 50% estarán cubiertos mientras que los demás no lo estarán. Además, esta práctica habitual resulta en este caso ser la que maximiza la proporción de observaciones correctamente clasificadas, es decir, aquellas en que se predice correctamente la situación de cobertura. Es así que la cobertura se predice usando como línea de corte una probabilidad de 0.50 de estar cubierto.

En el Cuadro 15 se clasifican las observaciones de acuerdo a la situación real y predicha de cobertura.

Cuadro 15

*Cobertura real y predicha en la muestra reducida**
Área Metropolitana, 1992.

	Verdaderos		
Predichos	con cobertura	sin cobertura	Total
con cobertura	1448	345	1793
sin cobertura	158	392	550
Total	1606	737	2343

Pr(prededir c/cob. verdadero c/cob.)	90.16%
Pr(prededir s/cob. verdadero s/cob.)	53.19%
Pr(verdadero c/cob. predicho c/cob.)	80.76%
Pr(verdadero s/cob. predicho s/cob.)	71.27%

Pr(prededir c/cob. verdadero s/cob.)	46.81%
Pr(prededir s/cob. verdadero c/cob.)	9.84%
Pr(verdadero s/cob. predicho c/cob.)	19.24%
Pr(verdadero c/cob. predicho s/cob.)	28.73%

Clasificados correctamente	78.53%
----------------------------	--------

* Se predice "con cobertura" si la probabilidad estimada es mayor o igual a 0.5

En las columnas se separan los individuos cubiertos de los no cubiertos, mientras que en las filas se agrupan de acuerdo a si se predijo o no cobertura. Como puede observarse, un 78.5% de las observaciones está correctamente clasificada. Este modelo tiene la capacidad de detectar el 90% de los casos en que se observa cobertura y el 53% en los que no se observa. En cuanto a su valor predictivo, la probabilidad de que realmente un individuo esté cubierto cuando se predijo que así lo estaba es del 81%, mientras que la probabilidad de no estarlo cuando se predijo que el individuo no tenía seguro es del 71%. El mayor problema que aparece es una alta probabilidad de predecir cobertura cuando en realidad no la hay (46.8%).¹⁷ Esto lleva a cierta sobreestimación de la proporción de gente con cobertura: mientras que la proporción predicha es del 76.5% en la muestra reducida, el verdadero porcentaje es del 68.5%.

La información anterior puede verse de una forma algo más detallada en el Cuadro 16. Para la confección del mismo se ordenaron las 2343 observaciones usadas en la estimación en forma creciente de acuerdo a la probabilidad estimada. Las mismas se separaron en 10 grupos (en la última columna se reporta el número de observaciones en cada grupo). La probabilidad estimada promedio de cada grupo se encuentra en la segunda columna. En las columnas 3 y 4 aparecen los totales de gente, por grupo, con cobertura observada y estimada, respectivamente. Las columnas 5 y 6 presentan cifras similares pero para los no cubiertos. Como puede observarse, si bien para ciertos grupos se subestima y

¹⁷ Este error podría reducirse sólo a costa de aumentar otro tipo de error (la probabilidad de no predecir cobertura cuando sí la hay). Ver el Apéndice.

para otros se sobrestima la cobertura, en el agregado termina habiendo sobreestimación.

Cuadro 16

*Cobertura real y predicha en la muestra reducida.
Área Metropolitana, 1992.*

Grupo	Probabilidad	Con Cobertura		Sin Cobertura		Total
		Observado	Estimado	Observado	Estimado	
1	0.2671	36	39.1	199	195.9	235
2	0.4441	88	83.8	146	150.2	234
3	0.5879	118	122	116	112	234
4	0.6943	162	151	73	84	235
5	0.7686	165	171.7	69	62.3	234
6	0.8184	190	185.5	44	48.5	234
7	0.8679	190	197.9	45	37.1	235
8	0.9147	210	208.6	24	25.4	234
9	0.9535	216	218.9	18	15.1	234
10	0.9928	231	227.6	3	6.4	234

4.5. Extensión de las predicciones para la muestra completa

Una vez estimadas las probabilidades para la muestra reducida pueden extenderse las estimaciones al resto de las observaciones. Como se explicó antes, a los Jubilados y Pensionados se les asigna su probabilidad *ex ante* de estar cubiertos (98.2%), mientras que para los Cónyuges y menores de 18 años se usa el Teorema de Bayes para inferirla a partir de las probabilidades estimadas para los Jefes de Hogar. Debido a datos incompletos y al problema de inconsistencia en el reporte de los ingresos, se termina trabajando con 7618 observaciones (3220 menos que la base original). A partir de ahora se hace referencia a esta muestra como la *muestra completa*. Al igual que antes, se considera que un individuo tiene cobertura si la probabilidad estimada es superior a 0.5.

En el Cuadro 17 se clasifican nuevamente las observaciones, ahora para la muestra completa, de acuerdo a su situación de cobertura observada y predicha. Un 82% de las observaciones está bien clasificada (un porcentaje algo mayor que el de la muestra reducida). Se detecta un 93% de los casos con cobertura y un 52% de los sin cobertura. La proporción de gente con cobertura en la muestra completa es del 73.5% y se estima un 81.21%.

Cuadro 17

*Cobertura real y predicha en la muestra completa.*¹⁸

¹⁸ La muestra completa cuenta con 7638 observaciones. Sin embargo la parte B del Cuadro 17 sólo reporta información acerca de 7618 debido a datos faltantes sobre cobertura real de 20 individuos.

Área Metropolitana, 1992.

A) Predicciones*

	obs.	%
con cobertura	6203	81.21
sin cobertura	1435	18.79
Total	7638	100

B) Cobertura real y predicha

	Verdaderos		
Predichos*	con cobertura	sin cobertura	Total
con cobertura	5211	976	6187
sin cobertura	392	1039	1431
Total	5603	2015	7618

Pr(prededir c/cob. verdadero c/cob.)	93.00%
Pr(prededir s/cob. verdadero s/cob.)	51.56%
Pr(verdadero c/cob. predicho c/cob.)	84.22%
Pr(verdadero s/cob. predicho s/cob.)	72.61%

Pr(prededir c/cob. verdadero s/cob.)	48.44%
Pr(prededir s/cob. verdadero c/cob.)	7.00%
Pr(verdadero s/cob. predicho c/cob.)	15.78%
Pr(verdadero c/cob. predicho s/cob.)	27.39%

Clasificados correctamente	82.04%
----------------------------	--------

* Se predice "con cobertura" si la probabilidad estimada es mayor o igual a 0.5

En gran medida, el error de predicción se introduce al fijar arbitrariamente un valor límite a la probabilidad para determinar cuando una persona está o no cubierta. Afortunadamente, el uso de este tipo de arbitrariedades puede evitarse cuando se pasa del análisis muestral al de la población.

5. Predicción de la proporción de población con cobertura

5.1. Predicciones para 1992

Como se explicó antes, es posible traducir a cifras poblacionales conclusiones basadas en las encuestas de la EPH. Esto se hace mediante el uso de ponderadores que miden la "representatividad" de cada hogar de la encuesta y que se reportan en la misma. La muestra original representa a la población total del Área Metropolitana de Buenos Aires. Debido a que la muestra completa con la que se trabaja aquí no incluye a una parte de las observaciones, no es posible extender las conclusiones a toda la población sino solamente a

aquella parte representada por la muestra completa. De aquí en más se hará referencia a esa población como la *población representada*. Mientras que la población total es de 10.995.123, la población representada es de 7.751.915. Solamente sobre este último total pueden extenderse las conclusiones de la estimación.

Lo que surge del análisis es que el 74.08% de la población representada tiene cobertura, un total de 5.742.885 personas. El modelo predice cobertura para 5.768.775 personas, el 74.42% de la población representada. Es evidente lo bien que se comporta este modelo al momento de predecir la población con cobertura sobre el total representado.

Si bien no es posible realizar predicciones en relación a la población total, es posible hacer ciertas inferencias que permiten extender los resultados. Según la información con que se cuenta, la proporción de población no representada cubierta es significativamente inferior a la de la población representada. Esto es indicativo de algún componente sistemático en el reporte inconsistente de ingresos. Como se ve en el Cuadro 18, este porcentaje no alcanza al 60%.¹⁹

Cuadro 18

*Cobertura de la población no representada por la muestra completa.
Área Metropolitana, 1992.*

	obs.	%
con cobertura	1907257	58.81
sin cobertura	1294263	39.91
ns/nc	41688	1.29
Total	3243208	100

En el Cuadro 19 se muestra que casi un 70% de la población del Área Metropolitana tenía cobertura en 1992. Este porcentaje es algo inferior al 74.52% predicho para la población representada. Esta diferencia se debe, obviamente, al bajo índice de cobertura de la población no representada.

Pese a este problema, las predicciones realizadas siguen siendo muy informativas y útiles. En primer lugar, el error de predicción resulta ser bastante acotado. En segundo lugar, puede tomarse al 74.42% como una estimación de máxima, es decir, interpretarlo como un límite superior al porcentaje de población con cobertura.

¹⁹ Otra particularidad que se observa es un índice de respuestas ns/nc significativamente superior a la media. Este es otro elemento que puede estar relacionado con cierta sistematicidad en el reporte inconsistente de los ingresos.

Cuadro 19

*Cobertura de la población total.
Área Metropolitana, 1992.*

	obs.	%
con cobertura	7650142	69.58
sin cobertura	3283391	29.86
ns/nc	61590	0.56
Total	10995123	100

5.2. Predicciones para 1996

Como ya se explicó antes, el módulo de educación y servicios de salud de la EPH sólo se incluyó en mayo de 1992 para algunos aglomerados. Por este motivo, las sucesivas encuestas no cuentan con información sobre cobertura.

En esta sección se predice la población con cobertura para el Área Metropolitana en 1996, usando los parámetros estimados en secciones previas. Al igual que para la base de datos de 1992, para este año hay problemas debido a la inconsistencia en el reporte de los ingresos, por lo que para la predicción se pierdan observaciones. A pesar de esto, y como se vio en la parte final de la sección anterior, pueden interpretarse las predicciones como límites superiores del índice de cobertura de la población. Cuanto mayor sea la población representada en relación a la población total, más cercano estará este límite superior a la verdadera proporción de cobertura poblacional, *ceteris paribus*. Mientras que para el Área Metropolitana en 1992 la población representada es un 70% de la total el porcentaje aumenta significativamente para 1996 alcanzando al 85%.

Para 1996 la EPH de este área cuenta con un total de 11.398 observaciones representando a una población de 11.561.653 habitantes. La muestra completa cuenta con 9.726 observaciones que representan a una población de 9.868.540 habitantes.

Según las estimaciones, un 72.29% de la población representada tendría cobertura (7.134.403 habitantes).

La comparación de este resultado con el obtenido para 1992 resulta útil para reconocer el papel que juega la proporción de la población que está representada. Como ya fue señalado, una menor proporción de población representada genera un sesgo a sobreestimar el nivel de cobertura de la población. Si se comparara directamente la proporción de cobertura estimada del 74.4% en 1992 con el 72.3% en 1996 sin tener en cuenta este sesgo, se inferiría que el nivel de cobertura de la población cayó entre esos años. Una comparación válida implicaría considerar a ambas estimaciones como estimaciones de máxima de la proporción de población cubierta y reconocer que la primera

de ellas es susceptible a una mayor sobreestimación (la población representada en proporción a la total aumentó significativamente en 1996). Así, mientras que la reducción estimada del índice de cobertura es del orden de los dos puntos porcentuales, hay un aumento considerable (del 70 al 85%) en la proporción de la población que está representada. Teniendo en mente estas consideraciones no parece factible que la diferencia en las estimaciones se deba a una real reducción en el nivel de cobertura de la población sino más bien a una diferencia de sesgos. Efectivamente, comparando el 72.3% estimado para 1996 con el nivel de cobertura observado para 1992 del 69.6%, es posible pensar que la proporción de población cubierta no varió significativamente entre un año y otro. Si se considera constante la proporción de la población no representada con cobertura (58.8%), el nivel de cobertura estimado sobre el total de la población sería del 70.32%, prácticamente el mismo que el observado en 1992. El Cuadro 20 resume los resultados de esta sección.

Cuadro 20

Predicciones de la proporción de población cubierta.

Área metropolitana. Años 1992 y 1996.

	Población Total	Población Representada	Población Rep./ Población Total	Pob. con Cobertura*/ Pob. Representada
1992	10.995.123	7.751.915	70.50%	74.42%
1996	11.561.653	9.868.540	85.36%	72.29%

* estimada

PARTE II: LA DETERMINACIÓN DE LA FORMA DE COBERTURA

La Parte I del trabajo se concentró en analizar de qué forma las distintas características demográficas y socioeconómicas de las personas afectan su propensión a contar con un seguro de salud. Se consideró que para cada individuo había dos alternativas, cobertura o no cobertura, y que la probabilidad de una u otra venía determinada por una serie de características personales.

Esta parte del trabajo se concentra en el análisis de la determinación del tipo de seguro de salud. En particular, se estudia cómo distintos atributos de la población determinan distintos tipos de cobertura. El análisis que se lleva a cabo es condicional a tener cobertura, es decir, se hace para aquellas personas que se observan aseguradas. El ejercicio empírico consiste en estimar un modelo de la probabilidad de tener cada uno de los cuatro tipos de cobertura: obligatoria y a cargo de un familiar (OAC o estado 1), obligatoria en carácter de titular (OT o estado 2), voluntaria y a cargo de un familiar (VAC o estado 3) y voluntaria como titular (VT o estado 4).

6. El método empírico

Todos los individuos que poseen cobertura en salud son susceptibles de ser caracterizados por uno de los cuatro tipos de cobertura mencionados (afiliado obligatorio a cargo de un familiar, afiliado obligatorio titular del seguro, afiliado voluntario a cargo de un familiar o afiliado voluntario titular de la cobertura).

En la Sección 2, punto 2.2, se presentó el argumento de los modelos de índices lineales. Decíamos que, en este contexto, la situación de cobertura que se observa para un individuo es aquella para la cual es más propenso dadas sus características (las propias y las de su familia). El mismo argumento se aplica a la forma de cobertura observada.

Al igual que en la primera parte, se considera que el índice lineal Xb viene determinado por ciertas características socioeconómicas del individuo y de su hogar. Designemos por Y a una variable aleatoria que indica el tipo de seguro observado, donde $Y=1, 2, 3$ o 4 representa a OAC, OT, VAC y VT, respectivamente. Según el modelo logístico multinomial, las probabilidades para estas cuatro formas de cobertura vienen dadas por:

$$Prob(Y_i=j|X_i) = \frac{e^{X_i\beta_j}}{\sum_{k=1}^4 e^{X_i\beta_k}} \quad \text{donde } j= 1, 2, 3 \text{ o } 4$$

Así especificado, este modelo presenta una indeterminación en el sentido de que hay más de un posible valor de los $\mathbf{b}_j$ con $j=1, 2, 3, 4$ que genera la misma distribución de probabilidades para Y .²⁰ Para eliminar esta indeterminación se fija arbitrariamente $\mathbf{b}_1=0$.²¹ Las probabilidades resultantes después de esta normalización son:

$$Prob(Y_i=j|X_i) = \frac{e^{X_i\beta_j}}{1 + \sum_{k=2}^4 e^{X_i\beta_k}} \quad \text{para } j = 2, 3, 4$$

$$Prob(Y_i=1|X_i) = \frac{1}{1 + \sum_{k=2}^4 e^{X_i\beta_k}} \quad \text{para } j = 1$$

Al fijar arbitrariamente $\mathbf{b}_1$ en cero, OAC, o el estado 1, se constituye en la categoría base. La única consecuencia que tiene esto es condicionar la forma en que deben interpretarse los coeficientes estimados para los otros tipos de cobertura. Así, el índice lineal $X_i\mathbf{b}_j$ mide cuánto más propenso es el individuo i a tener el seguro j que un seguro OAC. Por su parte $\mathbf{b}_{jk}$ es el efecto marginal de la característica x_k sobre el índice lineal. Pero debido a la no linealidad de las probabilidades en los coeficientes $\mathbf{b}$, éstos no pueden interpretarse como los efectos marginales de las distintas variables sobre las probabilidades. Además, a causa de que hay múltiples ecuaciones (una para cada tipo de cobertura, excepto la categoría base), tampoco es posible hacer inferencias en base al signo de los coeficientes. Por ejemplo, supongamos que una de las variables independientes de nuestro modelo toma solamente los valores 0 y 1, y que estamos interesados en evaluar el efecto de esta variable sobre la probabilidad de cada una de las posibles formas de cobertura. Para ser más explícitos, consideremos la variable *sexo*, que toma el valor uno para los hombres y cero para las mujeres. Supongamos que el coeficiente de esta variable para la segunda categoría, OT, es positivo. Lo que uno tendería a pensar es que los hombres tienen una mayor probabilidad de ser OT. Si bien generalmente este será el caso, podría darse que el coeficiente de la variable *sexo* para otra categoría sea aún mayor, haciendo que la probabilidad de ser OT caiga en relación a la de esa otra categoría para los hombres. Esto no sucede en los modelos de elección binaria como el analizado en la primera parte. Cuando hay solamente dos posibilidades, un aumento en la probabilidad de una se corresponde con una disminución idéntica en la probabilidad de la otra, por lo que el signo

²⁰ Notar que si $\mathbf{b}_j^* = \mathbf{b}_j + \mathbf{1}$, para cualquier vector $\mathbf{1}$, $\mathbf{b}_j^*$ y $\mathbf{b}_j$ dan lugar a las mismas probabilidades para cada uno de los cuatro posibles valores de Y . Esto se debe a que todos los términos que incluyen $\mathbf{1}$ se cancelan.

²¹ Con esta restricción, si $\mathbf{b}_j^* = \mathbf{b}_j + \mathbf{1}$, $\mathbf{b}_j^*$ y $\mathbf{b}_j$ dan lugar a distintas probabilidades para cualquier $\mathbf{1} \neq \mathbf{0}$.

de los coeficientes indica en qué dirección se mueven las probabilidades de elegir una u otra alternativa. Cuando las alternativas son múltiples, el signo sólo puede ser interpretado en relación a la categoría base. Que ser hombre aumente la probabilidad de elegir OT respecto de la categoría base no implica que la probabilidad de otra categoría no aumente aún más. Esto se da porque mientras en el numerador de la expresión para la probabilidad de ser OT aparecen sólo los coeficientes estimados para esa categoría, en el denominador intervienen los de todos los estados.

Para facilitar la interpretación de las estimaciones, el modelo logístico multinomial puede expresarse en términos de *relative odds ratios* (ROR).

Simplifiquemos la notación de la siguiente forma:

$$P_j = \text{Prob}(Y_i=j|X_i) \text{ para } j=1, 2, 3 \text{ y } 4$$

Luego,

$$P_j / P_n = \exp(X_i \mathbf{b}_j) / \exp(X_i \mathbf{b}_n) = \exp[X_i(\mathbf{b}_j - \mathbf{b}_n)]$$

donde X_i y $\mathbf{b}$ son vectores de $(1 \times S)$ y $(S \times 1)$, respectivamente. En particular, y ya que $\mathbf{b}_1 = 0$.

$$P_j / P_1 = \exp(X_i \mathbf{b}_j) / \exp(X_i \mathbf{b}_1) = \exp[X_i(\mathbf{b}_j - \mathbf{b}_1)] = \exp(X_i \mathbf{b}_j)$$

Consideremos un cambio en una unidad de la variable x_k . Denotemos como P_j' a la probabilidad del estado j dado el vector $X_i' = X_i + \phi$, donde ϕ es un vector de las mismas dimensiones que X_i con un uno en la k -ésima posición y todos los demás elementos iguales a cero. Es decir,

$$P_j' = \text{Prob}(Y_i=j|X_i') \text{ para } j=1, 2, 3 \text{ y } 4 \text{ y } X_i' = X_i + \phi$$

Así,

$$P_j' / P_1' = \exp(X_i' \mathbf{b}_j) / \exp(X_i' \mathbf{b}_1) = \exp[X_i'(\mathbf{b}_j - \mathbf{b}_1)] = \exp(X_i' \mathbf{b}_j)$$

Finalmente, tenemos que:

$$(P_j' / P_1') / (P_j / P_1) = \exp(X_i' \mathbf{b}_j) / \exp(X_i \mathbf{b}_j) = \exp[(X_i' - X_i) \mathbf{b}_j] = \exp(\phi \mathbf{b}_j) = \exp(\mathbf{b}_{jk})$$

donde $\mathbf{b}_{jk}$ es el coeficiente estimado para el estado j de la variable x_k .

La última expresión es el *relative odds ratio* o ROR para el estado j y un cambio de una unidad en la variable x_k . Claramente estos cocientes son siempre positivos. Cuando $\mathbf{b}_{jk}$ es negativo, el ROR está entre cero y uno, indicando que un aumento unitario de la variable

x_k reduce la probabilidad de tener el tipo de seguro j en relación al tipo OAC. A un b_{jk} positivo le corresponde un ROR mayor que uno. Por ejemplo, si $b_{jk}=0.7$, el ROR correspondiente es de 2.01, indicando que cuando x_k aumenta en una unidad se duplica la relación entre las probabilidades del seguro j y el 1

7. Estimación del modelo

7.1. Los datos

Nuevamente se usa la EPH del Área Metropolitana de mayo de 1992. En la muestra original hay 7450 individuos con cobertura de salud (ver Cuadro 1). Una vez eliminados todos aquellos que hacen reportes incoherentes de ingresos y otros para los que la información sobre ciertas variables relevantes es incompleta, la muestra se reduce a 4891 observaciones.

7.2. Las variables

La variable dependiente del modelo logístico multinomial es Y , y representa el tipo de cobertura. $Y=1$ si OAC, $Y=2$ si OT, $Y=3$ si VAC e $Y=4$ si VT. Dado que se fijó OAC como la categoría base, se estiman tres ecuaciones que corresponden a las otras tres.

La única variable explicativa que se agrega en este modelo en relación a las que se usaron para el modelo logit de la primera parte es $edad2$. Esta variable es el cuadrado de la edad y se incluye para permitir una influencia no lineal de ésta sobre el índice lineal Xb .²²

En el Cuadro 21 se presentan los principales estadísticos descriptivos de las variables incluidas en el modelo, desagregados por tipo de cobertura.

²² En el modelo logit de la primera parte no se incluyó la edad al cuadrado ya que en estimaciones preliminares y bajo diferentes especificaciones, el coeficiente de esta variable nunca resultó significativamente distinto de cero.

Cuadro 21

*Estadísticos descriptivos de las variables por tipo de cobertura.
Área Metropolitana, 1992.*

	OAC		OT		VAC		VT		Totales	
Variable	Media	D. Std.	Media	D. Std.	Media	D. Std.	Media	D. Std.	Media	D. Std.
edad	25.92	20.03	50.26	18.67	26.55	19.42	39.26	16.76	37.14	22.51
sexo	0.32	0.47	0.62	0.49	0.35	0.48	0.65	0.48	0.47	0.50
casado	0.38	0.49	0.59	0.49	0.37	0.48	0.65	0.48	0.49	0.50
jefe	0.04	0.20	0.67	0.47	0.11	0.31	0.65	0.48	0.35	0.48
conyuge	0.35	0.48	0.14	0.35	0.31	0.46	0.16	0.37	0.25	0.43
hijo	0.57	0.50	0.12	0.33	0.53	0.50	0.16	0.36	0.35	0.48
otrael	0.04	0.19	0.07	0.26	0.05	0.22	0.04	0.19	0.05	0.23
menores	0.53	1.11	0.73	1.21	0.51	1.01	0.94	1.15	0.63	1.15
jubilado	0.03	0.16	0.35	0.48	0.03	0.16	0.05	0.23	0.16	0.37
empleo	0.14	0.35	0.65	0.48	0.22	0.42	0.74	0.44	0.40	0.49
liaee	5.57	0.68	5.70	0.68	6.00	0.74	6.14	0.76	5.70	0.71
asal	0.09	0.29	0.59	0.49	0.13	0.33	0.31	0.46	0.32	0.47
fjub	0.03	0.16	0.36	0.48	0.03	0.17	0.05	0.23	0.17	0.38
cuentap	0.04	0.20	0.01	0.11	0.06	0.24	0.27	0.45	0.04	0.21
otrasf	0.01	0.11	0.03	0.16	0.05	0.21	0.17	0.38	0.03	0.17
noperci	0.83	0.38	0.00	0.07	0.74	0.44	0.19	0.39	0.43	0.49
preesc	0.07	0.25	0.00	0.02	0.08	0.28	0.02	0.15	0.04	0.19
primi	0.36	0.48	0.16	0.37	0.26	0.44	0.08	0.27	0.25	0.43
primc	0.20	0.40	0.37	0.48	0.15	0.36	0.18	0.38	0.26	0.44
seci	0.23	0.42	0.15	0.36	0.22	0.42	0.20	0.40	0.19	0.39
secc	0.08	0.27	0.16	0.37	0.13	0.33	0.19	0.40	0.12	0.33
univi	0.04	0.20	0.07	0.25	0.10	0.30	0.16	0.37	0.07	0.25
univc	0.03	0.17	0.09	0.29	0.06	0.25	0.16	0.37	0.07	0.25
obs.	1962		2096		571		262		4891	
% total	40.11		42.85		11.67		5.36		100	

7.3. La estimación del modelo

La estimación del modelo logístico multinomial se hace sobre las 4891 observaciones descriptas en el Cuadro 21. Se usa la rutina de Stata V, cuya salida de regresión se presenta en el Cuadro 22. Como ya fue señalado, la categoría omitida es OAC, por lo que tanto los coeficientes estimados como los ROR deben interpretarse en relación a la misma.

Cuadro 22

Estimaciones del Modelo Logístico Multinomial. Categoría base: OAC.
Área Metropolitana, 1992.

Multinomial regression									Number of obs = 4891			
									chi2(63) = 5319.04			
									Prob > chi2 = 0.0000			
Log Likelihood= -2901.9094									Pseudo R2 = 0.4782			
	OT				VAC				VT			
Y	Coef.	D.Std.	PV	ROR	Coef.	D.Std.	PV	ROR	Coef.	D.Std.	PV	ROR
edad	0.1809	0.0216	0.0000	1.1983	0.0064	0.0198	0.7470	1.0064	0.0838	0.0299	0.0050	1.0875
edad2	-0.0016	0.0002	0.0000	0.9984	0.0000	0.0002	0.9390	1.0000	-0.0008	0.0003	0.0130	0.9992
sexo	0.9674	0.1763	0.0000	2.6310	0.0607	0.1228	0.6210	1.0625	0.5386	0.2288	0.0190	1.7137
casado	-0.3077	0.1991	0.1220	0.7351	-0.5315	0.2266	0.0190	0.5877	-0.2315	0.2635	0.3800	0.7933
conyuge	-1.9073	0.2288	0.0000	0.1485	-0.6775	0.2494	0.0070	0.5079	-2.2697	0.3117	0.0000	0.1033
hijo	-1.1912	0.2610	0.0000	0.3039	-0.8386	0.3064	0.0060	0.4323	-2.1770	0.3572	0.0000	0.1134
otrael	-0.4635	0.3061	0.1300	0.6291	-0.5881	0.3429	0.0860	0.5554	-1.2710	0.4327	0.0030	0.2806
menores	-0.1232	0.0683	0.0710	0.8841	0.0306	0.0678	0.6520	1.0310	-0.1025	0.0888	0.2490	0.9026
jubilado	3.1420	0.6405	0.0000	23.1496	-0.4042	0.6789	0.5520	0.6675	1.0427	0.7030	0.1380	2.8369
empleo	2.6383	0.5232	0.0000	13.9892	0.4150	0.3786	0.2730	1.5143	1.7605	0.4566	0.0000	5.8155
liaee	-0.2546	0.1090	0.0190	0.7752	0.7706	0.0773	0.0000	2.1610	0.3611	0.1283	0.0050	1.4349
fjub	-0.2547	0.4673	0.5860	0.7751	0.6807	0.6098	0.2640	1.9753	0.1774	0.5903	0.7640	1.1941
cuentap	-3.7768	0.2783	0.0000	0.0229	0.0572	0.2605	0.8260	1.0589	0.0550	0.2504	0.8260	1.0566
otrasf	-1.6663	0.3963	0.0000	0.1889	0.9171	0.4146	0.0270	2.5020	1.1302	0.4074	0.0060	3.0961
noperci	-3.3028	0.5919	0.0000	0.0368	0.6450	0.4029	0.1090	1.9059	0.5093	0.4922	0.3010	1.6641
preesc	-1.1263	2.6305	0.6690	0.3242	-0.1985	0.3681	0.5900	0.8200	0.1629	0.6240	0.7940	1.1769
primi	-0.9937	0.3035	0.0010	0.3702	-0.6381	0.2877	0.0270	0.5283	-1.3259	0.3886	0.0010	0.2656
primc	-1.0105	0.2583	0.0000	0.3640	-0.4800	0.2665	0.0720	0.6188	-1.2594	0.3221	0.0000	0.2838
seci	-0.9416	0.2632	0.0000	0.3900	-0.4024	0.2669	0.1320	0.6687	-0.8967	0.3163	0.0050	0.4079
secc	-0.4994	0.2538	0.0490	0.6069	-0.0311	0.2656	0.9070	0.9694	-0.5733	0.3083	0.0630	0.5637
univi	-0.4028	0.2941	0.1710	0.6685	0.1642	0.2945	0.5770	1.1785	-0.2080	0.3424	0.5440	0.8122
constante	-1.8526	1.0028	0.0650		-5.1843	0.8251	0.0000		-4.4054	1.1619	0.0000	

sombreado si la variable es significativa al 5%.

El estadístico Chi2(63)=5319.04 presentado en la parte superior derecha del Cuadro 22 es significativo a menos del 1%, indicando que el modelo es relevante para explicar las distintas formas de cobertura. En el primer panel del Cuadro 23 se presentan tests de significatividad conjunta de las variables, uno para cada ecuación. El resultado es que en los tres casos se rechaza la hipótesis nula de que los coeficientes sean simultáneamente iguales a cero.

El sexo es relevante en la determinación del tipo de cobertura.²³ En particular, los hombres son más propensos a ser titulares del seguro (OT o VT).²⁴ El ROR correspondiente

²³ Cada vez que se diga que un coeficiente es significativamente distinto de cero, se tendrá como referencia un nivel de significatividad del 5% a menos que se indique lo contrario.

²⁴ Salvo que se indique lo contrario, siempre se usa como base de la comparación una cobertura obligatoria y

a OT es de 2.6. Esto indica que el hecho de ser hombre casi triplica la relación entre las probabilidades de OT y OAC, respecto a la que corresponde a una mujer con sus mismos atributos. Para un seguro tipo VT esta relación es de 1.7. Además, y según los resultados reportados en el cuarto panel del Cuadro 23, no puede rechazarse la hipótesis nula de que los efectos del sexo sobre la probabilidad de OT y VT son iguales. Por otra parte, el sexo no es significativo para explicar una diferencia de probabilidades entre las dos alternativas que involucran afiliación a cargo de un familiar (VAC y OAC). De lo anterior se puede inferir que, aún después de tener en cuenta los efectos del rol que cada individuo desempeña en el hogar (relación de parentesco), su situación laboral, su estado civil, etc., los hombres son más propensos a ser titulares del seguro que las mujeres. Si bien según el resultado encontrado en la primera parte del trabajo el sexo no era relevante para explicar la situación de cobertura, sí contribuye en la determinación del tipo de seguro.

Pese a que ser casado aumenta las chances de ser OAC sobre VAC, no puede rechazarse la hipótesis nula de que los coeficientes de casado para los tres estados son simultáneamente iguales a cero (ver segundo panel del Cuadro 23).²⁵ Esto último indicaría que diferencias en el estado civil no explican diferencias en las probabilidades de los distintos tipos de cobertura, pese a que el estado civil sí es un determinante significativo del hecho de tener cobertura, como se vio en la primera parte del trabajo.

La relación de parentesco omitida es Jefes de Hogar. Distintas relaciones de parentesco dan lugar a diferencias significativas en las probabilidades de cada tipo de cobertura. Los coeficientes negativos para todas las dummies de parentesco indican que los jefes de hogar son los que tienen menos chances de estar a cargo en un seguro obligatorio. En términos de *relative odds ratios*, la relación entre la probabilidad de OAC y la de VT es 7 veces mayor para un cónyuge y más del triple para los hijos de lo que es para el jefe del hogar. La misma relación pero entre OAC y VAC es de más de 2 para ambos grupos. En resumen, tanto los cónyuges como los hijos son más propensos a tener un tipo de cobertura obligatoria a cargo que otra forma de seguro. Por su parte, los jefes de hogar resultan más propensos a ser titulares del seguro o bien a tener un seguro voluntario. En la primera parte del trabajo se encontró cierta evidencia acerca de la relación entre la cobertura de los jefes del hogar y la de los cónyuges y menores de 18 años. Estos últimos resultados de alguna manera confirman que, por lo general, los jefes determinan la cobertura “arrastrando” a los demás miembros de la familia.

Cuadro 23

Tests sobre los coeficientes del modelo multinomial.

a cargo de un familiar (OAC).

²⁵ Sí puede rechazarse esta hipótesis nula para un nivel de significatividad del 10%.

Tests de Wald

Ho: todos los coeficientes son cero excepto la constante

OT	chi2(21) = 748.80	Prob > chi2 = 0.0000
VAC	chi2(21) = 204.14	Prob > chi2 = 0.0000
VT	chi2(21) = 497.79	Prob > chi2 = 0.0000

Ho: los coeficientes son simultáneamente cero en todas las ecuaciones

casado	chi2(3) = 6.03	Prob > chi2 = 0.1103
menores	chi2(3) = 4.52	Prob > chi2 = 0.2105
jubilado	chi2(3) = 33.26	Prob > chi2 = 0.0000
fjub	chi2(3) = 4.41	Prob > chi2 = 0.2206
cuentap	chi2(3) = 249.71	Prob > chi2 = 0.0000
otrasf	chi2(3) = 106.56	Prob > chi2 = 0.0000
noperci	chi2(3) = 44.26	Prob > chi2 = 0.0000
preesc	chi2(3) = 0.62	Prob > chi2 = 0.8928
secc	chi2(3) = 5.69	Prob > chi2 = 0.1276
univi	chi2(3) = 3.53	Prob > chi2 = 0.3167
edad edad2	chi2(6) = 84.26	Prob > chi2 = 0.0000
fjub cuentap otrasf noperci	chi2(12) = 352.65	Prob > chi2 = 0.0000
preesc primi primc seci secc univi	chi2(18) = 49.52	Prob > chi2 = 0.0001

Ho: los coeficientes son simultáneamente iguales a cero en la ecuación VAC

edad edad2	chi2(2) = 0.83	Prob > chi2 = 0.6614
fjub cuentap otrasf noperci	chi2(4) = 5.42	Prob > chi2 = 0.2466
preesc primi primc seci secc univi	chi2(6) = 20.89	Prob > chi2 = 0.0019

Ho: los coeficientes son iguales para las categorías OT y VT

edad edad2	chi2(2) = 21.13	Prob > chi2 = 0.0000
sexo	chi2(1) = 3.36	Prob > chi2 = 0.0669
liaee	chi2(1) = 21.82	Prob > chi2 = 0.0000
conyuge hijo otrarel	chi2(3) = 10.37	Prob > chi2 = 0.0157
empleo	chi2(1) = 2.35	Prob > chi2 = 0.1254
preesc primi primc seci secc univi	chi2(6) = 3.70	Prob > chi2 = 0.7174

Ho: los coeficientes son iguales para las categorías OT y VAC

edad edad2	chi2(2) = 55.75	Prob > chi2 = 0.0000
sexo	chi2(1) = 22.45	Prob > chi2 = 0.0000
liaee	chi2(1) = 75.35	Prob > chi2 = 0.0000
conyuge hijo otrarel	chi2(3) = 23.48	Prob > chi2 = 0.0000
empleo	chi2(1) = 16.42	Prob > chi2 = 0.0001
preesc primi primc seci secc univi	chi2(6) = 4.91	Prob > chi2 = 0.5560

sombreado si no se puede rechazar la hipótesis nula a un nivel de significatividad del 5%.

El número de hijos menores no explica las diferencias en las formas de cobertura observadas entre individuos. Más aún, según el segundo panel del Cuadro 23 ni siquiera

puede rechazarse la hipótesis nula de que sus coeficientes sean simultáneamente iguales a cero para las tres ecuaciones. Según el modelo logit de la primera parte, la cantidad de hijos menores de 18 años tampoco era relevante para explicar la probabilidad de cobertura.

La variable jubilado, como ya fue descrita en la primera parte de presente trabajo, toma el valor uno para todos aquellos que reportándose inactivos, dicen ser jubilados o pensionados. La relación entre las probabilidades de OT y OAC es 23 veces mayor para un jubilado que para alguien que no lo es. En cuanto a las demás categorías, esta variable no explica las diferencias observadas en el tipo de seguro.

Tener empleo aumenta las chances de ser titular de un seguro en relación a las de tener una cobertura OAC. Como muestra el Cuadro 23, no hay una diferencia significativa en los efectos que tiene esta variable sobre las probabilidades de los estados que involucran titularidad del seguro, ni sobre las formas de cobertura a cargo. Por lo tanto, puede decirse que tener empleo contribuye a determinar el hecho de ser titular del seguro o afiliado a cargo de un familiar, aumentando las chances de la primera alternativa sobre la segunda.

El ingreso ajustado por adulto equivalente y economías de escala internas al hogar es significativo en las tres ecuaciones. Un mayor ingreso por mes, aumenta las chances de cualquiera de las coberturas voluntarias sobre OAC, siendo los efectos correspondientes a VT y VAC significativamente distintos. Curiosamente, un aumento del ingreso reduce la probabilidad de OT relativa a OAC. En la primera parte del trabajo se encontró que aumentos en el ingreso conllevan a aumentos en la probabilidad de tener cobertura. Según estos últimos resultados, aumentos en el ingreso de alguien asegurado también aumentan las chances de contratar seguros voluntarios.

Con respecto a las fuentes de ingresos, cabe aclarar nuevamente que las categorías se construyeron a partir de un reagrupamiento sobre las presentadas por el INDEC en la base de usuarios de la EPH. Así, *fjub* por ejemplo, incluye a los que perciben solamente jubilaciones y pensiones o combinaciones de estas y otras fuentes de ingresos. Por este motivo las variables *jubilado* y *fjub* no son idénticas. El Cuadro 22 presenta evidencia que para un mismo nivel de ingresos la probabilidad de OT relativa a OAC es significativamente afectada por la fuente de procedencia de los mismos. La categoría omitida es la correspondiente a ingresos derivados del trabajo asalariado. Según los resultados, no hay diferencias significativas entre los que perciben jubilaciones y/o pensiones y los asalariados. Ser cuentapropista, no percibir ingresos o hacerlo de fuentes tales como rentas y beneficios aumenta las chances de ser afiliado a cargo de un seguro obligatorio por sobre uno obligatorio como titular. En cuanto a las dos categorías que involucran voluntariedad de la afiliación, las variables que representan a las fuentes de

ingresos no son conjuntamente significativas.²⁶ Vinculando estos resultados con los de la primera parte del trabajo, los cuentapropistas son los que tienen menos probabilidad de tener cobertura y de tenerla, son los más propensos a ser OAC en lugar de OT.

El máximo nivel educativo alcanzado es relevante en la determinación de la forma de cobertura. La categoría omitida es *univc*. Los coeficientes de *univi* no son significativos en ninguna de las tres ecuaciones, lo que implica que este nivel no se diferencia del universitario completo en su efecto sobre las chances de uno u otro estado respecto del estado 1. Curiosamente lo mismo ocurre con el coeficiente de *preesc*. No se puede rechazar la hipótesis nula de que los coeficientes de *secc* sean simultáneamente iguales a cero en todas las ecuaciones. Lo mismo para los de *preesc* y *univi*. Además, según los dos últimos paneles del Cuadro 23, los efectos de este grupo de variables son significativamente iguales para todas las alternativas.²⁷ Comparando OT con OAC, todas las categorías educacionales, menos *preesc* y *univi*, son relevantes. Los ROR son menores que uno, lo que indica que en relación a los que tienen un nivel universitario completo, pertenecer a estos grupos aumenta las chances de ser OAC sobre OT. Además, los coeficientes de *primi*, *primc* y *seci* de la ecuación OT son significativamente iguales entre sí y significativamente menores que el de *secc*.²⁸ Esto implica que es posible distinguir tres grupos educativos: 1) educación superior o universitario completa o incompleta, 2) educación secundaria completa y 3) educación secundaria incompleta y primaria completa o no. Algo similar ocurre en relación a VT. La diferencia es que en este caso el efecto de *secc* tampoco es significativamente distinto del de *univc*, por lo que se diferencian sólo dos grupos según niveles educativos. Entonces, cuanto mayor el nivel educativo, mayores las chances de ser titular de un seguro relativas a las de tener un seguro obligatorio a cargo. Sobre la base de esto y de las conclusiones del modelo de elección binaria de la Parte I, hay evidencia de que la educación incrementa la probabilidad de tener un seguro de salud, y podría esperarse encontrar a los asegurados más educados contratando un seguro voluntario o bien siendo titulares de la cobertura.

Al incluir edad y edad al cuadrado como variables explicativas, es posible captar efectos no lineales del “envejecimiento” sobre el índice lineal. Como se señaló antes, $X_i b_j$

²⁶ No es posible rechazar la hipótesis nula que los cuatro coeficientes son simultáneamente cero para la ecuación VT. El $\text{Chi}^2(4)=8.63$. Los resultados del mismo test pero con relación a la ecuación VAC se reportan en el Cuadro 23.

²⁷ Los tests reportados en el Cuadro 23 muestran que los coeficientes son iguales entre pares de ecuaciones. El estadístico de un test de que son simultáneamente iguales para las tres ecuaciones es $\text{Chi}^2(12)=11.39$, al que corresponde un p-value de 0.49.

²⁸ Este test no está reportado en el Cuadro 23. El estadístico para evaluar la hipótesis nula de que los coeficientes de *primi*, *primc* y *seci* son simultáneamente iguales en la ecuación OT es $\text{Chi}^2(2)=0.13$, con un p-value de 0.98. El correspondiente para evaluar la misma hipótesis pero agregando *secc* es $\text{Chi}^2(3)=7.53$, que es significativo al 5.6%.

debería interpretarse como la propensión del individuo i a asegurarse bajo la alternativa j en relación a la base.

El efecto marginal de la edad sobre el índice lineal del estado j para el individuo i viene dado por:

$$\frac{\partial(X_i b_j)}{\partial \text{edad}} = b_{j,\text{edad}} + b_{j,\text{edad}^2} \text{edad}_i$$

Un valor positivo para esta expresión indicaría que un aumento en la edad aumenta la propensión a tener un seguro j en relación a la de tener uno OAC.

Según las estimaciones, la edad no modifica las probabilidades relativas entre las dos formas de cobertura a cargo, mientras que sí lo hace entre éstas y las alternativas que involucran titularidad. Los coeficientes de edad para OT y VT son ambos positivos, mientras que los de edad al cuadrado son negativos (todos estos coeficientes son significativamente distintos de cero). Esto último indica que los índices lineales de OT y VT son estrictamente cóncavos respecto a la edad, es decir, la propensión a ser titular aumenta hasta cierta edad para luego disminuir a medida que la persona envejece. Para explorar más en detalle los efectos de la edad sobre las probabilidades de una u otra alternativa, es útil concentrar el análisis en grupos de individuos con características particulares. Para esto se procede a un ejercicio de simulación que asigna a cada grupo un determinado perfil. Por supuesto que es posible encontrar una enorme cantidad de combinaciones. A fines descriptivos se trabajará con cuatro grupos solamente. El análisis se hace basándose en los gráficos 3 al 6.

Las curvas del Gráfico 3 miden $X_i b_j$ para $j = \text{OT}$ e $i = G1, G2, G3$ y $G4$. En el Gráfico 4 se presenta lo mismo pero para $j = \text{VT}$. El primer grupo de individuos ($i = G1$) representa a hombres jefes de familia y casados, cuya fuente de ingresos es el trabajo asalariado y tienen educación universitaria o superior completa. Consideremos a $G1$ como el grupo base, ya que los demás surgen de ir cambiando una característica por vez partiendo de este. $G2$ representa a hombres con los mismos atributos que $G1$ pero que son cónyuges en lugar de jefes de hogar. Los cuentapropistas conforman $G3$, mientras que $G4$ reúne a hombres con un nivel educativo máximo de primaria completa. Como resultado del signo de los coeficientes y como ya se mencionó, todas las curvas son cóncavas. Como la pendiente de estas curvas mide el efecto marginal de la edad sobre $X_i b_j$, la concavidad implica que este efecto decrece con la edad.

Gráfico 3

Índice lineal de OT.

Simulación en base a las estimaciones reportadas en el Cuadro 22.

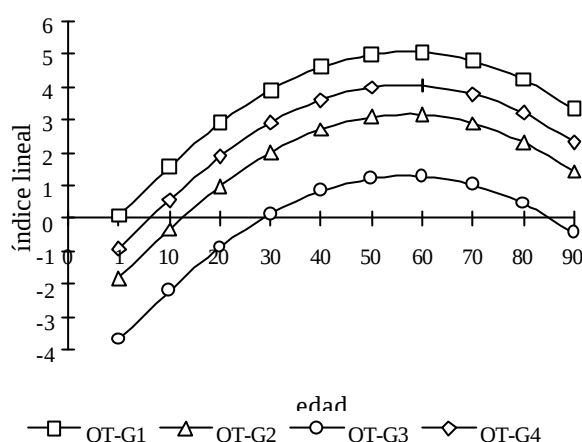
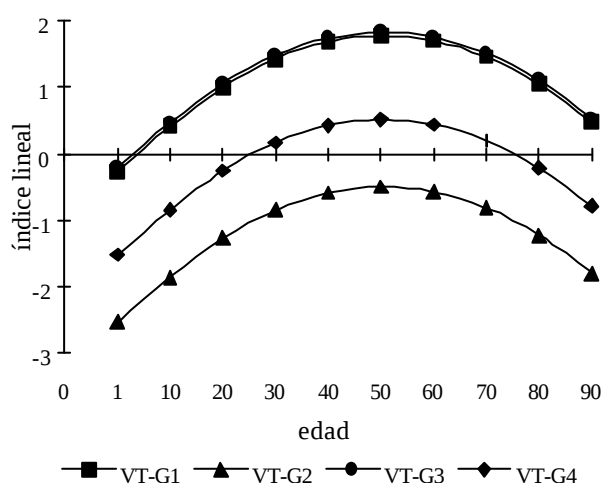


Gráfico 4

Índice lineal de VT.

Simulación en base a las estimaciones reportadas en el Cuadro 22.



G1: jefes de hogar sexo masculino y casados, asalariados con educación universitaria completa.
 G2: cónyuges sexo masculino y casados, asalariados con educación universitaria completa.
 G3: jefes de hogar sexo masculino y casados, cuentapropistas con educación universitaria completa.
 G4: jefes de hogar sexo masculino y casados, asalariados con educación primaria completa.
 Las variables continuas se evalúan en las medias.

Concentrémonos primero en el Gráfico 3. Cada una de las curvas corresponde a uno de los grupos definidos antes. La primer curva empezando desde arriba es para el grupo base (G1). La segunda corresponde al grupo 4 (los de “educación primaria”). La curva correspondiente a G2 (“los cónyuges”) es la tercera y la última curva corresponde a G3 (“los cuentapropistas”). Toda la curva del grupo base está por encima del eje de abscisas, lo

que indica que esta gente es siempre más propensa a ser OT que OAC. Para G4 y G2 (segunda y tercera curva, respectivamente) puede decirse lo mismo.²⁹ La gente del grupo de los cuentapropistas (última curva) tiene una mayor propensión a ser OAC que OT hasta una edad aproximada de 30 años. A partir de esa edad se revierte el orden entre estas dos alternativas.

En relación al Gráfico 4 puede hacerse un análisis similar. Aquí lo que se evalúa es la propensión a tener un seguro VT comparada con la de uno OAC. Como puede verse, las dos curvas superiores prácticamente se superponen y están siempre sobre el eje de abscisas. Estas son las correspondientes a los grupos 1 y 3 (grupo base y cuentapropistas, respectivamente), que son más propensos a tener un seguro VT que uno OAC.³⁰ Por la ubicación de la curva correspondiente a G4 se entiende que a mayor educación, mayor propensión a tener un seguro VT. Sólo después de los 25 años una persona de este grupo tendría más chances de contratar un seguro voluntario en carácter de titular que estar afiliado a cargo de un seguro obligatorio. Por último, los “cónyuges” (G2) siempre tienen más probabilidad de tener un seguro obligatorio a cargo que una cobertura voluntaria en carácter de titular.

En los gráficos 5 y 6 se reproducen algunas de las curvas de los gráficos 3 y 4. Cabe recordar aquí que la edad no genera una diferencia de probabilidades significativa entre VAC y OAC (ver Cuadro 22), por lo que la curva que mide Xb_{VAC} aparece como una línea horizontal. El Gráfico 5 se construye para el grupo 3, es decir “los cuentapropistas”, para quienes se ve claramente que, independientemente de la edad, VT es el tipo de seguro más probable. Esto es consistente con el análisis hecho más arriba: de todas las relaciones de parentesco, los jefes de hogar son los más propensos a ser titulares del seguro o bien a contratar un seguro voluntario, y de este grupo son los cuentapropistas los más propensos a ser titulares de la cobertura. Con relación a las tres alternativas restantes, se puede decir que para cualquier edad este tipo de individuos siempre es más propenso a ser OAC que VAC, y que para los individuos mayores de 30 años la segunda alternativa más probable es OT. Para los mayores de 40 años el orden de los tipos de cobertura desde más a menos probable es VT, OT, OAC y VAC.

Gráfico 5

Índice lineal de los distintos tipos de cobertura.

Jefes de hogar cuentapropistas con educación universitaria completa.

²⁹ Esta afirmación se refiere al rango relevante de edades que surge de considerar los demás atributos que están siendo tenidos en cuenta para delimitar los grupos.

³⁰ En el Cuadro 22 se presenta evidencia de que el coeficiente de *cónyuge* no es significativamente distinto del de *asal*, que es la categoría omitida.

Simulación en base a las estimaciones reportadas en el Cuadro 22.

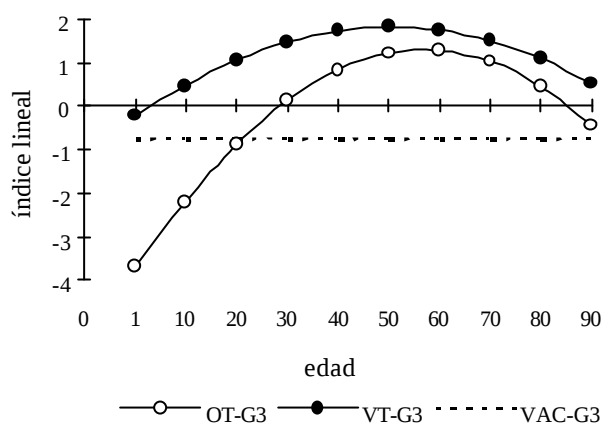
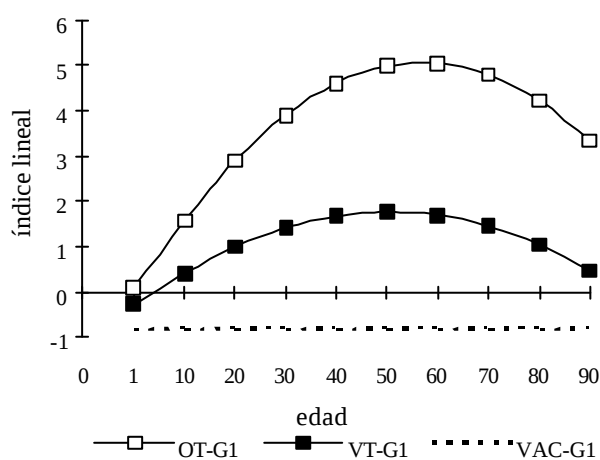


Gráfico 6

Índice lineal de los distintos tipos de cobertura.

Jefes de hogar asalariados con educación universitaria completa.

Simulación en base a las estimaciones reportadas en el Cuadro 22.



El Gráfico 6 es similar al 5 pero se construyó para el grupo base, es decir, jefes de hogar de sexo masculino, casados, asalariados y con educación universitaria o superior completa. Para este grupo el ordenamiento de las alternativas de acuerdo a su probabilidad no cambia con la edad del individuo. De más a menos probable el orden es OT, VT, OAC y VAC.

Aprovechando que aquí este orden no se revierte, se procede a analizar con más detalle el efecto de la edad sobre cada tipo de cobertura. Las mayores chances de tener un

seguro como titular sobre uno OAC se acentúan progresivamente a medida que la edad aumenta, hasta llegar a un máximo y luego disminuyen. La máxima propensión a tener OT la tiene un individuo de 57 años de edad. Para VT esto se da a los 50 años. Esto significa que sería mucho más probable observar la elección de OT para hombres del grupo 1 con edades rondando los 50 que para otros más jóvenes.

7.4. Los efectos marginales

Hasta aquí se ha descripto cómo las distintas variables contribuyen a explicar la probabilidad de cada alternativa de cobertura. También se exploró la influencia de la edad sobre las chances relativas entre los distintos tipos de seguro. Usando los gráficos anteriores también es posible ver los efectos marginales de la edad sobre los índices lineales. En particular, se evidencian efectos marginales mayores en las curvas correspondientes a OT respecto de las que corresponden a VT, implicando que para el grupo 1 las probabilidades de OT se acentúan a medida que aumenta la edad (Gráfico 6). Con relación al grupo 3, estos mayores efectos marginales no alcanzan para hacer que los cuentapropistas tengan más chances de elegir OT sobre VT (Gráfico 5).

De los gráficos anteriores también puede verse claramente el efecto marginal sobre el índice lineal de cambios en los atributos, que no son otra cosa que los coeficientes estimados. Tanto en el Gráfico 3 como en el 4 las curvas son paralelas entre sí. Esto se debe a que por la especificación del modelo las distintas características no interactúan con la edad, haciendo que los efectos de las distintas dummies sobre el índice lineal sean constantes. Estos efectos se miden como las distancias verticales entre las curvas. Del gráfico 3 surge que el efecto de aumentar la educación desde un nivel primario a uno universitario o superior puede medirse como la distancia vertical entre las dos curvas más altas. Este efecto es menor que el de pasar de ser cónyuge a jefe de hogar, que a su vez es menor que el de pasar de ser cuentapropista a asalariado. La diferencia con el Gráfico 4 es que en este último pasar de ser asalariado a cuentapropista aumenta la propensión de ser VT relativa a la de ser OAC.

Como en el caso del modelo logit de la primera parte, para obtener los efectos marginales de las distintas variables sobre las probabilidades de contratar cada tipo de seguro es necesario hacer algunos cálculos.

El efecto sobre la probabilidad del estado j de un cambio marginal en la variable x_k viene dado por:

$$\frac{\partial P_j}{\partial x_k} = P_j [\beta_{jk} - \sum_{i=1}^4 P_i \beta_{ik}]$$

Como puede apreciarse, el signo y magnitud de esta expresión depende de las magnitudes relativas de los coeficientes y de la probabilidad de cada una de las alternativas de cobertura. Esta expresión sirve para evaluar los efectos marginales de las variables continuas solamente. Una ventaja de examinar los efectos marginales es que, a diferencia de los coeficientes, estos no dependen de la categoría base elegida. Por este motivo, es posible computar los efectos marginales también para esta. Lamentablemente, el programa de computación utilizado no reporta los efectos marginales correspondientes a la categoría base, por lo que sólo se analizarán los correspondientes a las otras tres formas de cobertura.

En el Cuadro 24 se presentan los efectos sobre las probabilidades de OT, VAC y VT de un cambio marginal porcentual del ingreso ajustado. Los efectos marginales se calculan para cuatro grupos distintos de individuos, todos ellos jefes de hogar y casados, y se evalúan en las medias de las variables continuas. Los dos primeros grupos son de hombres, uno de asalariados y el otro de cuentapropistas, y los otros dos son de mujeres (también asalariadas y cuentapropistas, respectivamente).

Un aumento de un 1% en el ingreso ajustado tiene como consecuencia aumentar la probabilidad de contratar un seguro voluntario mientras que reduce las de tener uno OT para los cuatro grupos. Ante un aumento del ingreso, las mujeres cuentapropistas son las que experimentan el mayor aumento en la probabilidad de contratar VAC, mientras que el mayor aumento en la probabilidad de contratar VT se produce para los hombres cuentapropistas. Estos últimos, a su vez, son los que sufren una mayor caída en la probabilidad de tener un seguro obligatorio titular. Quienes experimentan la menor caída en la probabilidad de tener este tipo de cobertura son los hombres asalariados.

Discriminando los efectos según el nivel educativo máximo alcanzado, se ve que a mayor nivel educativo, menor es el efecto marginal sobre la probabilidad de tener VAC y mayor sobre la de tener VT. Así, el mayor efecto marginal positivo en relación a la cobertura VAC se registra para las mujeres cuentapropistas con educación primaria completa, en tanto que para VT resulta mayor el de los hombres cuentapropistas con educación universitaria o superior completa.

Cuadro 24

Efectos marginales del ingreso sobre la probabilidad de cada tipo de cobertura.

Jefes de hogar casados.

<i>A. Hombres</i>			
<i>A.1. Asalariados</i>			
	OT	VAC	VT
primc	-0.040	0.010	0.024
secc	-0.042	0.009	0.029
univc	-0.040	0.006	0.032
<i>A.2. Cuentapropistas</i>			
	OT	VAC	VT
primc	-0.086	0.058	0.077
secc	-0.096	0.054	0.079
univc	-0.102	0.036	0.090
<i>B. Mujeres</i>			
<i>B.1. Asalariadas</i>			
	OT	VAC	VT
primc	-0.066	0.021	0.032
secc	-0.067	0.020	0.039
univc	-0.062	0.013	0.044
<i>B.2. Cuentapropistas</i>			
	OT	VAC	VT
primc	-0.050	0.076	0.056*
secc	-0.060	0.073	0.053*
univc	-0.069	0.052	0.064

* Significativos al 10%.

Con el Cuadro 25 es posible explorar cómo el cambio en determinados atributos de los individuos afecta a las probabilidades de contratar cada uno de los tipos de seguro, es decir, los efectos marginales de las variables binarias. Para su construcción se calcularon las medias de las probabilidades predichas por el modelo para distintas características de los individuos, tanto ajustadas como sin ajustar. Por característica sin ajustar se hace referencia a la que se observa en la muestra, mientras que las características ajustadas son producto de una simulación. Por ejemplo, en el primer panel se analiza el efecto del género de las personas sobre la probabilidad de las distintas formas de cobertura. En las dos primeras columnas se reportan las probabilidades medias predichas para los hombres y las mujeres en la muestra, es decir, los cálculos se hacen en base al género observado (datos sin ajustar). Para obtener las predicciones medias ajustadas se simula que todos los individuos de la muestra son primero mujeres, manteniendo constantes todas las demás características. Luego se predicen las probabilidades y se promedian. Así se obtienen las cifras de la columna 3. Para la columna 4 se procede de la misma forma pero suponiendo primero que todos los individuos son hombres. La diferencia entre estas dos columnas puede

interpretarse como el efecto del género sobre las probabilidades de los distintos tipos de cobertura, luego de haber controlado por todas las demás características de los individuos. El mismo procedimiento se aplica para el resto del cuadro, variando, en cada panel, la relación de parentesco, la fuente de ingresos y el máximo nivel educativo alcanzado, respectivamente.

Cuadro 25

Medias de las probabilidades predichas. Sin ajustar y ajustadas.

Predicciones y simulación en base al modelo del Cuadro 22.

	<i>Sin ajustar</i>		<i>Ajustados</i>	
	mujeres	hombres	mujeres	hombres
OAC	51.37	27.45	42.15	37.52
OT	30.78	56.43	40.02	45.65
VAC	14.25	8.77	12.38	11.13
VT	3.59	7.34	5.45	5.70
Obs.	2589	2302	4891	4891
	cónyuges	jefes	cónyuges	jefes
OAC	51.06	4.35	48.21	37.52
OT	30.80	82.82	38.67	45.65
VAC	13.65	3.15	10.22	11.13
VT	4.49	9.68	2.90	5.70
Obs.	1204	1705	4891	4891
	asal.	cuentap.	asal.	cuentap.
OAC	7.89	35.43	37.52	52.01
OT	83.99	14.41	45.65	11.78
VAC	3.25	15.38	11.13	17.61
VT	4.87	34.78	5.70	18.61
Obs.	1581	220	4891	4891
	primc.	univc.	primc.	univc.
OAC	27.06	12.95	43.86	37.52
OT	62.95	65.35	42.92	45.65
VAC	6.10	8.89	9.49	11.13
VT	3.90	12.81	3.73	5.70
Obs.	1295	333	4891	4891

Lo que se puede observar comparando las columnas sin ajustar con las ajustadas es que para estas últimas el efecto del cambio en la característica particular siempre es menor.³¹ Esto se debe a que en la muestra sin ajustar se suman otros efectos al del atributo analizado, haciendo que generalmente se sobrestime el impacto que el mismo tiene sobre

³¹ La única excepción es para el efecto en la probabilidad de contratar VAC de un cambio en la educación.

las probabilidades.

El primer panel del cuadro sugiere que pese a que un hombre y una mujer tengan la misma edad, ingresos, educación, etc., la mujer tiene mayor probabilidad de tener una cobertura a cargo. Algo similar sucede en el segundo panel. Luego de controlar por todos los demás atributos, los jefes de hogar resultan ser más propensos a tener cobertura en carácter de titular que los cónyuges. Según la muestra sin ajustar estos últimos tenían mayores chances de ser VAC. Sin embargo, luego de controlar por las demás características el efecto de la relación de parentesco va en la dirección contraria.

En relación a la fuente de ingresos, los cuentapropistas tienen mayores probabilidades de ser voluntarios o afiliados a cargo, mientras que las mayores chances de ser OT las tienen los asalariados. Por su parte, el panel 4 deja ver que aquellos que tienen nivel universitario o superior completo son más propensos a ser voluntarios o afiliados titulares, mientras que para un nivel educativo máximo de escuela primaria completa es más probable ser OAC. Seguramente esto se deba a que el nivel educativo tiene alguna correlación con las características del empleo que no son captadas por la fuente de ingresos.

7.5. Poder predictivo del modelo

El primer panel del Cuadro 26 es una tabla de clasificación que compara las predicciones del modelo con el tipo de cobertura observada. La regla de predicción es que un individuo tiene cobertura j si $P_j = \max\{P_i \text{ para } i=1, 2, 3 \text{ y } 4\}$.³² En los restantes paneles del cuadro se sintetizan algunos de los resultados que surgen de la tabla del primer panel.

En primer lugar, casi un 79% de las observaciones están bien clasificadas. Las probabilidades de detectar OT y OAC verdaderos son altas, mientras que sólo un 5.6% de los VAC verdaderos son predichos como tales.

Como señala *Arias (1996)*, es posible definir un *pseudo error tipo I* y un *pseudo error tipo II* para cada uno de las cuatro formas de cobertura de la forma siguiente:

Error tipo I: predecir que un individuo no tiene el seguro j cuando sí lo tiene.

Error tipo II: predecir que un individuo tiene el seguro j cuando no lo tiene.

En el último panel del cuadro se presentan estos errores para los cuatro seguros posibles. El error tipo I es elevado para las dos alternativas voluntarias, siendo especialmente alto en el caso de VAC. El error tipo II se mantiene bastante bajo para todas las categorías, siendo particularmente chico para VAC y VT. Esto lleva a una gran subestimación de la proporción de individuos con formas de afiliación voluntaria.

³² Esta regla es equivalente a predecir cobertura en el modelo binomial si la probabilidad estimada es mayor o igual a 0.5.

Cuadro 26

Tipo de cobertura observada y predicha.

Se predice cobertura tipo j si $P_j = \max\{P_i \text{ para } i=1, 2, 3 \text{ y } 4\}$

Predichos	Verdaderos				Total
	OAC	OT	VAC	VT	
OAC	1710	30	424	61	2225
	76.85	1.35	19.06	2.74	100
	87.16	1.43	74.26	23.28	45.49
OT	209	2033	91	109	2442
	8.56	83.25	3.73	4.46	100
	10.65	96.99	15.94	41.6	49.93
VAC	21	4	32	5	62
	33.87	6.45	51.61	8.06	100
	1.07	0.19	5.6	1.91	1.27
VT	22	29	24	87	162
	13.58	17.9	14.81	53.7	100
	1.12	1.38	4.2	33.21	3.31
Total	1962	2096	571	262	4891
	40.11	42.85	11.67	5.36	100
	100	100	100	100	100

Pr(prededir OAC verdadero OAC)	87.16%
Pr(prededir OT verdadero OT)	96.99%
Pr(prededir VAC verdadero VAC)	5.60%
Pr(prededir VT verdadero VT)	33.21%
Pr(verdadero OAC predicho OAC)	76.85%
Pr(verdadero OT predicho OT)	83.25%
Pr(verdadero VAC predicho VAC)	51.61%
Pr(verdadero VT predicho VT)	53.70%

Clasificados correctamente	78.96%
----------------------------	--------

Pr(prededir no OAC verdadero OAC)	12.84%
Pr(prededir no OT verdadero OT)	3.01%
Pr(prededir no VAC verdadero VAC)	94.40%
Pr(prededir no VT verdadero VT)	66.79%
Pr(prededir OAC verdadero no OAC)	17.58%
Pr(prededir OT verdadero no OT)	14.63%
Pr(prededir VAC verdadero no VAC)	0.69%
Pr(prededir VT verdadero no VT)	1.67%

De todos modos, y como sugiere *Greene (1997)*, no hay realmente buenas medidas de bondad de ajuste para los modelos de elección binaria y multinomial. De hecho, usando la regla de predicción mencionada, en muestras desbalanceadas es muy usual predecir mayor número de unos para las categorías con la mayor frecuencia en la muestra. Este es

justamente el caso de las dos categorías obligatorias.

PARTE III: CONCLUSIONES

Como se señaló en la introducción, en este trabajo se presentaron dos avances. Por un lado se estudió de qué manera se determina la cobertura en salud de los individuos, utilizando información de la Encuesta Permanente de Hogares del Area Metropolitana de Buenos Aires de mayo de 1992. Por el otro, se propuso un método estadístico que permite predecir el nivel de cobertura de la población utilizando información limitada.

En la primera parte del trabajo se estudió la determinación de la situación de cobertura de un individuo (tener o no tener seguro), mientras que la determinación de las distintas formas que ésta puede tomar se analizó en la Parte II. Ambos análisis se llevaron a cabo explorando la relación que existe entre la cobertura en salud y distintas características socioeconómicas y demográficas de las personas, buscando identificar el perfil de atributos que caracteriza a cada una de las formas de seguro. En primer lugar se usó un modelo logit para estimar la probabilidad de tener un seguro de salud. Para el análisis del tipo de seguro, se distinguieron cuatro posibles alternativas: seguro obligatorio a cargo de un familiar, seguro obligatorio en carácter de titular, seguro voluntario a cargo y seguro voluntario como titular. A este problema se lo consideró condicional a tener un seguro, y se usó para su estimación un modelo logístico multinomial. Los resultados de las estimaciones se resumen a continuación.

a) El sexo no influye en la probabilidad de tener un seguro. Sin embargo, es más factible que un hombre con cobertura sea afiliado titular que a cargo que una mujer. Por su parte, mientras que estar casado aumenta la probabilidad de tener cobertura, no influye en la determinación de la forma de afiliación.

b) Ser Jefe de Hogar aumenta la probabilidad de contratar un seguro respecto a las otras relaciones de parentesco, y de tener seguros voluntarios o en carácter de titular. En ningún caso el número de hijos menores de 18 años contribuye en la determinación de la cobertura.

c) Un aumento en el ingreso ajustado aumenta la probabilidad de estar asegurado. Para alguien cubierto, más ingreso implica más chances de contratar seguros voluntarios. Independientemente del nivel de ingresos, la fuente de procedencia de los mismos influye no sólo sobre el hecho de tener cobertura sino también sobre el de qué tipo de seguro tener. En particular, los cuentapropistas son los que tienen menor probabilidad de estar asegurados, y de estarlo, son los más propensos a contratar seguros voluntarios o a ser OAC.

d) Según las estimaciones de la primera parte del trabajo, a mayor nivel educativo alcanzado, mayores las probabilidades de estar cubierto. En la Parte II se encontró que

cuanto menor es la educación, mayores son las chances de tener un seguro obligatorio y estar a cargo de un familiar que pertenecer a cualquiera de las otras tres categorías.

e) La edad del individuo es significativa en los dos modelos. El envejecimiento tiene un efecto positivo sobre la probabilidad de contratar un seguro de salud. En cuanto a las chances de ser titular de la cobertura versus tener un seguro OAC, la edad ejerce un efecto no lineal, ampliándose la brecha a medida que el individuo envejece.

El aporte metodológico hecho en este trabajo permitió predecir el nivel de cobertura de la población del Área Metropolitana para el año 1996, pese a no contar con información sobre la misma para ese año. Usando la información sobre cobertura que brinda la EPH de mayo de 1992, se estimó la probabilidad de tener cobertura y, a partir de dicha estimación, se predijo el nivel de cobertura de la población de dicho área para 1992. De la comparación de estas predicciones con los valores de cobertura observados, se concluyó que el modelo tiene una alta capacidad de predecir el verdadero valor del nivel de cobertura de la población representada. Mientras que la proporción de población representada con cobertura era del 74.08%, el modelo predijo un 74.42%. Luego se predijo que el nivel de cobertura para el Área Metropolitana en 1996 fue del 72.29%.

Como se explicó en su momento, el problema del reporte inconsistente que hacen muchos individuos encuestados obliga a eliminar un número de observaciones bastante considerable. Esto hace que las predicciones se limiten a la parte representada de la población. Por esto se impone una interpretación más rigurosa de los resultados. Una posibilidad es tomar los porcentajes de cobertura como estimaciones de máxima de la verdadera cobertura en salud. Otras alternativas pueden resultar de futura investigación en pos de identificar alguna sistematicidad en los reportes inconsistentes de ingresos. Esto podría ayudar a comprender más lo que sucede con el sector de la población representado por la parte de la muestra que esta vez fue necesario omitir.

APÉNDICE: Los errores de predicción del modelo de la probabilidad de cobertura

Para pasar de la estimación de la probabilidad a la predicción de uno de los dos posibles resultados (cubierto o no cubierto) es necesario establecer un valor límite de la probabilidad a partir del cual se considera que la persona en cuestión tiene seguro de salud. Como ya se explicó, la determinación de ese valor es arbitraria y usualmente se lo establece en 0.5, que para el modelo en cuestión es también el que maximiza la proporción de aciertos. Pese a esto, hay una tendencia a sobreestimar la cobertura: la probabilidad de predecir cobertura aunque no la haya (error tipo II) es del 46.81%. Como se señaló en un pie de página, sería posible reducir este error, aunque no sin costos. Para distintos valores del límite de la probabilidad estimada de cobertura, corresponden diferentes proporciones de aciertos y distintas probabilidades de error. En la tabla a continuación se presentan los resultados del ejercicio de aumentar progresivamente el valor del límite.

	Valor límite de la probabilidad de cobertura estimada						
	0.2	0.3	0.4	0.5	0.6	0.674	0.8
% de aciertos	73.28%	76.48%	77.55%	78.53%	77.64%	74.35%	66.03%
Error tipo I*	82.23%	68.52%	58.21%	46.81%	36.09%	28.36%	13.84%
Error tipo II*	1.25%	2.86%	6.04%	9.84%	16.06%	24.41%	43.21%

* El error tipo I se define como la probabilidad de predecir cobertura cuando en realidad no la hay. Por su parte, el error tipo II es la probabilidad de predecir no cobertura cuando sí la hay.

Como se observa, a medida que aumente el valor límite de la probabilidad el porcentaje de aciertos primero aumenta y luego disminuye. El valor máximo se registra cuando el límite es 0.5. El error tipo I siempre aumenta mientras que el error tipo II siempre disminuye. Esto da cuentas del *trade off* que hay entre uno y otro.

Partiendo de un límite de 0.5, que es el que se usa en este trabajo para predecir cobertura, reducir la probabilidad de predecir cobertura cuando no la hay implica no sólo aumentar la probabilidad de predecir no cobertura cuando sí la hay, sino también reducir la proporción de aciertos.

BIBLIOGRAFÍA

- Arias, Omar. "A Multinomial Choice Analysis of Work and Welfare Participation. *Mimeo. University of Illinois, Urbana, Champaign*. 1996.
- Bertranou, Fabio. "Health care services utilization and health insurance coverage: evidence from Argentina". *Mimeo*. 1998.
- Convenio (1998). "Acceso a la educación y la salud en la Provincia de Buenos Aires". *Convenio Ministerio de Economía de la Provincia de Buenos Aires - Facultad de Ciencias Económicas, UNLP*.
- Deaton, Angus. The Analysis of Household Surveys. A Microeconometric Approach to Development Policy. *Published for the World Bank*. 1997.
- Flood, M. Cristina V. de. "¿Cuánto gasta Argentina en Salud?" *VII Jornadas Internacionales de Economía de la Salud. VIII Jornadas Nacionales de Economía de la Salud de Argentina. "Innovaciones en la gestión de servicios de la salud"*. 1998.
- Gasparini, Leonardo. "La incidencia distributiva del sistema tributario en Argentina". En El sistema tributario en Argentina, *FIEL*. 1998.
- Greene, William H. Econometric Analysis. *Prentice Hall. Third Edition*. 1997.
- Gujarati, Damodar. Econometría. McGraw Hill. Segunda Edición.
- Long, J. Scott. Regression Models for Categorical and Limited Dependent Variables. *Advanced Quantitative Techniques in the Social Sciences. Volume 7. SAGE Publications*. 1997.
- Marquis, M. Susan and Long, Stephen H. "Worker demand for health insurance in the non-group market" *Journal of Health Economics* 14, 47-63, 1995.
- McFadden, D. Conditional logit analysis of qualitative choice behavior. En: Zarembka, P. (ed). *Frontiers in econometrics*. New York: Academic Press, 1974.
- Paulin, Geoffrey D. "Health Insurance Coverage for Low-Income Families: Findings from the Consumer Expenditure Survey". *Advancing the Consumer Interest. Volume 8. Number 2 / Fall 1996*.
- Rhine, Sherrie L. W. and Ying Chu NG. "The effect of employment status on private health insurance coverage: 1977 and 1987". *Health Economics* 7: 63-79, 1998.