

# **Especificación y estimación de un modelo de ecuaciones simultáneas con variable dependiente censurada para explicar el ingreso de la pareja conyugal en aglomerados urbanos de Argentina**

María Teresa Blaconá<sup>\*</sup>, María del Carmen García<sup>\*</sup>,  
Javier Bussi<sup>\*\*</sup>, Nora Ventroni<sup>\*\*</sup>

## **Resumen**

Este trabajo estima en forma conjunta los ingresos laborales individuales de los integrantes de la pareja conyugal, en distintos aglomerados de Argentina. Se emplea el método de Nelson y Olson (1978) para la estimación de sistemas de ecuaciones con variables censuradas y se detallan algunas limitaciones al utilizar la estimación sugerida por Amemiya (1979). Se realiza un análisis estadístico de las propuestas existentes en la bibliografía sobre el tema mediante un estudio empírico. Se destacan los inconvenientes que se presentan por el alto porcentaje de censura y se encuentran resultados sobre la formación del ingreso que difieren entre aglomerados.

JEL: C34-D19

## **Summary**

The two individual labor incomes of a couple are jointly estimated for different urban areas in Argentina. The method proposed by Nelson and Olson (1978) for the estimation of a system of equations with limited dependent variables is used. The difficulties that arise with the method proposed by Amemiya (1979) are also discussed. A statistical analysis is performed through an empirical study, using methods suggested by different authors. The problems that are present due to high percentage of censorship are stated and results for the explanation of income that differ among the urban areas are shown.

JEL: C34-D19

e-mail: mblacona@agatha.unr.edu.ar

---

<sup>\*</sup> Instituto de Investigaciones Teóricas y Aplicadas de la Escuela de Estadística, Facultad de Ciencias Económicas y Estadística, Consejo de Investigaciones, Universidad Nacional de Rosario.

<sup>\*\*</sup> Instituto de Investigaciones Teóricas y Aplicadas de la Escuela de Estadística, Facultad de Ciencias Económicas y Estadística, Universidad Nacional de Rosario.

# **Especificación y estimación de un modelo de ecuaciones simultáneas con variable dependiente censurada para explicar el ingreso de la pareja conyugal en aglomerados urbanos de Argentina**

María Teresa Blaconá<sup>\*</sup>, María del Carmen García<sup>\*</sup>,  
Javier Bussi<sup>\*\*</sup>, Nora Ventroni<sup>\*\*</sup>

## **I.- Introducción**

La información relativa a los individuos que se releva mediante la Encuesta Permanente de Hogares (EPH), que lleva a cabo el Instituto Nacional de Estadísticas y Censos (INDEC), ofrece importantes ventajas para el análisis empírico de la conformación del ingreso de la pareja conyugal. Las variables demográficas registradas se pueden incorporar a ecuaciones estimables, conjuntamente con otras características propias de los individuos que posibilitan capturar la heterogeneidad de los mismos.

Sin embargo, los datos microeconómicos poseen ciertas características que complican la modelación econométrica y su estimación. En el caso que nos ocupa, la dificultad es el alto porcentaje de mujeres que no participan en el mercado laboral (entre un 50% y 60% de mujeres con ingreso de fuente laboral cero, según el aglomerado urbano que se analice). Por lo tanto, la variable dependiente es limitada, hecho que se debe tener en cuenta para obtener estimaciones consistentes.

Existen trabajos pioneros sobre modelos microeconómicos uniecuacionales con variable dependiente limitada (por ejemplo, Tobin, 1958; Amemiya, 1974 y Heckman, 1974). Posteriormente se desarrollaron modelos para sistemas simultáneos con información limitada y se propusieron una variedad de estimadores diferentes para los coeficientes de la ecuación, siendo el de Amemiya (1974) un referente obligado sobre el tema. El estimador sugerido por Nelson y Olson (1978) se obtiene sustituyendo las variables explicativas endógenas por los valores predichos, mediante la aplicación de máxima verosimilitud Tobit. Amemiya (1978) sugiere un estimador alternativo que es asintóticamente más eficiente por contemplar la correlación existente entre los disturbios de las ecuaciones y que se obtiene aplicando Tobit a la forma reducida de las ecuaciones de interés y luego volviendo a la forma estructural mediante el enfoque de los mínimos cuadrados generalizados.

La comparación de la eficiencia de varios métodos de estimación se ha llevado a cabo en contextos especiales de modelos para variables respuestas limitadas. Amemiya (1978,1979) y Lee (1993), entre otros, mostraron que los estimadores en dos etapas que usan valores predichos pueden ser ineficientes respecto a alguna versión del estimador mínimo cuadrático generalizado de Amemiya (AGLS).

En este trabajo se estima un sistema de ecuaciones de ingresos laborales individuales de los integrantes de la pareja conyugal, avanzando sobre la propuesta de Blaconá et al.(2001). Se emplea el método econométrico propuesto por Nelson y Olson para la estimación de sistemas de ecuaciones con variables censuradas y se detallan algunas limitaciones al utilizar la estimación según Amemiya (1979). El propósito de este trabajo es realizar un análisis estadístico a través de distintas propuestas existentes en la bibliografía para la especificación y estimación de modelos de sistemas de ecuaciones simultáneas con variables censuradas. Esto se realiza mediante un estudio empírico sobre la formación del ingreso de la pareja conyugal, en distintos aglomerados de Argentina. La aplicación se efectúa a partir de la

---

<sup>\*</sup> Instituto de Investigaciones Teóricas y Aplicadas de la Escuela de Estadística, Facultad de Ciencias Económicas y Estadística, Consejo de Investigaciones, Universidad Nacional de Rosario.

<sup>\*\*</sup> Instituto de Investigaciones Teóricas y Aplicadas de la Escuela de Estadística, Facultad de Ciencias Económicas y Estadística, Universidad Nacional de Rosario.

información sumistrada por la EPH para los aglomerados urbanos Capital Federal, Conurbano Bonaerense, Gran Córdoba y Gran Mendoza, correspondiente a la primera onda de 1998.

En la sección II se presenta una breve reseña sobre la especificación y estimación de modelos multiecuacionales con variables dependientes censuradas. En la sección III se definen las variables que se incorporan al modelo para explicar el ingreso de la pareja conyugal. En la sección IV se realiza una descripción resumida de los aglomerados en base a las variables definidas en la sección III. En la sección V se presentan las estimaciones de los modelos para los distintos aglomerados y por último, en la sección VI se discuten los resultados hallados.

## II.- Modelos multivariados con variable dependiente limitada

Se puede definir un modelo de ecuación única para variable dependiente limitada como

$$Y_t^* = \beta' X_t + u_t \quad (II.1)$$

y

$$Y_t = h(Y_t^*) \quad (II.2)$$

donde,  $X_t$  es un vector de variables exógenas;  $Y_t^*$  es una variable aleatoria latente (es decir, no directamente observable);  $Y_t$  es la variable dependiente observada;  $u_t$  es un disturbio no observado que se supone no correlacionado e independiente de  $X_t$ , con distribución  $N(0, \sigma^2)$ ,  $\beta$  es un vector de parámetros desconocidos; y  $h$  es una función conocida que especifica la estructura de quiebre de la variable dependiente  $Y_t$ . Por ejemplo para la estructura Tobit (Tobin, 1958) la forma de  $h(\cdot)$  es

$$h(Y_t^*) = \begin{cases} Y_t^* & \text{si } Y_t^* > c_t \\ c_t & \text{si } Y_t^* \leq c_t \end{cases} \quad (II.3)$$

El valor del quiebre,  $c_t$ , debe ser una constante observada o una variable exógena.

Para el caso de un sistema de ecuaciones simultáneas de dos variables endógenas con una de ellas censurada, por ejemplo, una variable no restringida y la segunda restringida a ser no negativa, Amemiya (1974) y otros autores posteriores proponen un modelo que se puede escribir como

$$Y_{1t} = \alpha_1 Y_{2t} + \beta_1' X_t + u_{1t} \quad (II.4)$$

$$Y_{2t} = \begin{cases} \alpha_2 Y_{1t} + \beta_2' X_t + u_{2t} & \text{si RHS} > 0 \\ 0 & \text{en otro caso} \end{cases} \quad (II.5)$$

donde, RHS (right-hand side) significa que el segundo miembro de la ecuación es positiva.

Las correspondientes ecuaciones reducidas son

$$\left. \begin{matrix} Y_{1t} = \pi_1' X_t + v_{1t} \\ Y_{2t} = \pi_2' X_t + v_{2t} \end{matrix} \right\} \text{ si } \pi_2' X_t + v_{2t} > 0 \quad (II.6)$$

$$\left. \begin{array}{l} Y_{1t} = \pi_1' X_t + v_{1t} \\ Y_{2t} = 0 \end{array} \right\} \text{ si } (1 - \gamma_1 \gamma_2)(\pi_2' X_t + v_{2t}) \leq 0 \quad (\text{II.7})$$

En forma matricial las ecuaciones II.4 y II.5 se pueden expresar como

$$\Gamma Y_t \geq B X_t + w_t; \quad \text{si } Y_t \geq 0 \quad (\text{II.8})$$

y

$$(\Gamma Y_t)_i = (\Gamma X_t + u_t)_i, \quad \text{toda vez que } Y_{it} > 0.$$

Amemiya (1979) observa que el problema que presenta el sistema (II.8) es que se deben restringir los coeficientes de las variables endógenas a  $(1 - \gamma_1 \gamma_2) > 0$  por problemas de coherencia. Otro problema es la existencia de una única solución del modelo. El modelo tiene una única solución sí y sólo sí cada menor principal de  $\Gamma$  es positivo, la condición suficiente es que  $\Gamma + \Gamma'$  sea definida positiva.

Dado lo complicado de la aplicación usual de máxima verosimilitud con información limitada (MVIL) a la función de verosimilitud normal bivariada truncada, el autor citado propone la estimación por el análogo a mínimos cuadrados indirecto. Estas dificultades son bien conocidas tanto en la literatura estadística como econométrica.

Posteriormente, Nelson y Olson (1978) introducen un modelo de ecuaciones simultáneas con variables endógenas latentes. Suponiendo el mismo tipo de variables que en el modelo (II.4) y (II.5) de Amemiya, estos autores proponen la siguiente forma estructural del modelo

$$Y_{1t} = \alpha_1 Y_{2t}^* + \beta_1' X_t + u_{1t} \quad (\text{II.9})$$

$$Y_{2t}^* = \alpha_2 Y_{1t} + \beta_2' X_t + u_{2t} \quad (\text{II.10a})$$

$$Y_{2t} = \begin{cases} Y_{2t}^* & \text{si } Y_{2t}^* > 0 \\ 0 & \text{en otro caso} \end{cases} \quad (\text{II.10b})$$

Suponen que  $Y_{1t}$  e  $Y_{2t}$  son variables aleatorias observables,  $Y_{2t}^*$  es una variable latente no observable,  $X_t$  matrices de constantes conocidas y todos los otros términos son los que se definen generalmente para modelos de ecuaciones simultáneas. La correspondiente forma reducida del modelo es

$$Y_{1t} = \pi_1' X_t + v_{1t} \quad (\text{II.11})$$

$$Y_{2t}^* = \pi_2' X_t + v_{2t} \quad (\text{II.12a})$$

$$Y_{2t} = \begin{cases} Y_{2t}^* & \text{si } Y_{2t}^* > 0 \\ 0 & \text{en otro caso} \end{cases} \quad (\text{II.12b})$$

La diferencia de este último modelo con el de Amemiya, es que en las ecuaciones estructurales, la variable endógena del segundo miembro aparece en forma latente no restringida. Este cambio produce modificaciones importantes sobre las implicaciones del modelo, entre las que se pueden destacar: i) la forma reducida del modelo no requiere ninguna restricción sobre los coeficientes de las variables endógenas, ii) la ecuación de la forma reducida para  $Y_{1t}$  (no restringida) no depende de si la variable  $Y_{2t}$  (restringida) toma el valor o está por encima del

punto de quiebre y iii) esta especificación permite que la magnitud de la variable latente  $Y_{2t}^*$  afecte directamente el valor de  $Y_{1t}$ .

Las modificaciones planteadas en los puntos ii) y iii) pueden producir interpretaciones distintas según el modelo que se emplee en el caso de los ingresos de la pareja conyugal. Si se considera que  $Y_{1t}$  es el logaritmo del ingreso de la fuente laboral del varón e  $Y_{2t}$  el logaritmo del ingreso de la fuente laboral de la mujer, existe un porcentaje muy grande de mujeres con punto de quiebre para la censura  $c_t=0$ , por lo tanto preferir  $Y_{2t}$  como variable explicativa, puede afectar el desempeño de  $Y_{1t}$ . Se puede esperar que el desempeño de  $Y_{1t}$  sea diferente cuando se usa  $Y_{2t}^*$ .

La estimación del modelo de ecuaciones (II.9)-(II-10) no es sencilla ni viable por el procedimiento de MVIL. La estimación de la forma reducida es directa y se pueden obtener estimadores consistentes. La ecuación (II.11) es un modelo lineal estándar y puede ser estimada por mínimos cuadrados ordinarios (MCO), mientras que la ecuación (II.12) es un modelo Tobit y puede ser estimado por máxima verosimilitud. En cambio, para estimar los parámetros de la forma estructural se pueden obtener, en ciertas circunstancias, soluciones múltiples y estimadores ineficientes.

Para solucionar este problema, Nelson y Olson proponen un procedimiento análogo a mínimos cuadrados en dos etapas. El mismo se puede expresar brevemente de la siguiente manera: en el primer paso se estiman los parámetros de la forma reducida para cada ecuación por separado, la primera por mínimos cuadrados, y la segunda por máxima verosimilitud Tobit. Las estimaciones se usan para crear las variables instrumentales  $\hat{Y}_{1t} = \pi_1' X_t$  e  $\hat{Y}_{2t}^* = \pi_2' X_t$ , las cuales son al menos asintóticamente no correlacionadas con los términos de los disturbios. Estos instrumentos se sustituyen por sus contrapartes en el segundo miembro de las ecuaciones (II.9) y (II.10a). Las ecuaciones resultantes son tratadas como una ecuación lineal simple y un modelo Tobit, respectivamente. Luego, MCO y el análisis Tobit proveen los elementos para estimar, en la segunda etapa, los coeficientes estructurales, aplicando a cada ecuación máxima verosimilitud.

En la primera etapa de la estimación se considera que las covariancias entre los disturbios son cero. Un estimador más eficiente se puede obtener considerando la covariancia no cero entre los disturbios. Amemiya (1979) considera en la segunda etapa la covariancia entre los residuos de las ecuaciones y utiliza un método equivalente a mínimos cuadrados generalizados (AGLS). Huang, Sloan y Adamache (1987) proponen aplicar el algoritmo de maximización de la esperanza (EM) (Dempster, Laird, y Rubin, 1977) para calcular la función de verosimilitud en el modelo reducido resultante de suponer que las variables endógenas se distribuyen como una normal bivariada truncada.

### III.- Determinación y descripción de las variables que intervienen en el modelo

En este estudio se desea modelar el logaritmo del ingreso de la fuente laboral de cada uno de los cónyuges, en función de ciertas variables de interés. A continuación se presenta una descripción de las mismas.

#### **Variables endógenas**

Se decide trabajar con el ingreso laboral neto de cada uno de los integrantes de la pareja conyugal, debido a que es la variable registrada en la EPH. No se trabaja con la variable ingreso por hora, como suele ser costumbre en trabajos sobre ingreso por dos motivos: i) se considera que dividir por el número de horas trabajadas, cuando el período de referencia de esta variable no es el mismo que para la variable ingreso puede incluir una nueva fuente de error en la variable en estudio, la cual ya tiene problemas de registro; ii) el salario neto puede estar influenciado en algunos casos por el subsidio conocido como salario familiar.

Se consideran en forma conjunta los ingresos de los dos integrantes de la pareja conyugal.

Pero se trabaja con las variables transformadas, que se denominan:

- $Y_1$ : **liflv** : logaritmo del ingreso de fuente laboral del cónyuge (varón).
- $Y_2$ : **liflm**: logaritmo del ingreso de fuente laboral de la cónyuge (mujer).

En los casos en que los ingresos presentan valores iguales a cero, no se aplica la transformación y permanecen como tales. La variable  $Y_2$  presenta una distribución asimétrica con datos censurados a la izquierda, correspondientes a ingresos cero.

La consideración conjunta de los ingresos de los dos integrantes de la pareja conyugal, mediante la estimación de un sistema de ecuaciones simultáneas, puede ayudar a atenuar el problema de endogeneidad, al tiempo que permite tener en cuenta las relaciones entre los ingresos de ambos.

Existen estudios econométricos que consideran algunas de las características del cónyuge en una ecuación de ingreso individual e introduce el ingreso como variable exógena, al igual que la educación o alguna otra característica. Este criterio fue objetado tempranamente por Griliches(1977), en relación con la educación y por Boulier y Rosenzweig (1984) en lo que se refiere a las características llevadas por la persona al mercado matrimonial y tratadas como si fueran exógenas. Lam (1988) resaltó que una implicación empírica de cualquier modelo de ordenamiento marital es que las características de los esposos son endógenas al mercado matrimonial. Sin embargo, Jacobsen y Rayack(1996) no encontraron diferencias sustanciales entre los ingresos de los maridos independientemente del hecho que sus esposas trabajen o no.

### **Variables exógenas**

Las variables registradas en la EPH y que son utilizadas en la literatura económica para explicar el ingreso de fuente laboral (o su transformación logarítmica) denominadas variables exógenas (X), se definen para cada uno de los cónyuges. Para el caso de los varones se agrega a cada variable el sufijo v y para las mujeres se agrega el sufijo m.

Siguiendo la tradicional teoría del capital humano (Mincer, 1970), es indudable que las principales dificultades inherentes a su contraste empírico se encuentran en identificar las habilidades de las personas, entendidas como características que pueden ser consideradas dotaciones, y distinguirlas de las de inversión. Se definen las variables que pueden ser consideradas de inversión.

- **edad**: variable continua calculada a partir de la fecha de nacimiento declarada por el encuestado.

Esta variable se toma como *proxi* de la experiencia de la persona.

- **edad2**: variable edad elevada al cuadrado.

Se incluye para captar el comportamiento no lineal del retorno por año de actividad, que se supone creciente hasta una cierta edad y luego puede comenzar a decrecer lentamente.

- **medesco**: años de escolaridad declarados por el encuestado.

Típica variable del capital humano que mide el retorno por cada año adicional de educación del individuo. Otro aspecto a tener en cuenta es la relación entre los logros educativos de uno de los miembros de la pareja conyugal y los ingresos del otro miembro. Becker (1973) propuso una aplicación fundada en la teoría acerca de la formación de parejas basadas en la asociación de características. Sugirió que la educación de la esposa es una variable *proxy* de características no observadas que afectan la productividad del cónyuge en las actividades fuera del mercado.

En Estados Unidos se probó la existencia de una relación positiva entre ambas variables (Becker, 1983) y se interpreta como el resultado de la influencia de la educación en la calidad de flujos de servicios de apoyo que las esposas proporcionan a los maridos, y que incrementa la productividad de las actividades de mercado de éstos, en el marco de una fuerte división sexual de trabajo. Akerloff (1997) señaló, por otra parte, que la educación es una variable que

no sólo tiene importancia económica intrínseca, sino también significado social pues revela la posición social relativa de las personas, pudiendo ser éste un factor atractivo por sí mismo.

- **lifnl**: logaritmo del ingreso de fuente no laboral. Variable continua que es el logaritmo de la suma de todos los ingresos no laborales declarados por el encuestado. Si es cero, permanece como tal.

Se cuenta con cada una de las siguientes variables exógenas relativas al hogar, que pueden influenciar el ingreso de ambos cónyuges.

- **cont6**: cantidad de hijos menores de 6 años.
- **cont18**: cantidad de hijos entre 6 y 18 años.

En la mayoría de las parejas, suele ser el hombre el que recibe un subsidio por lo hijo (salario familiar) y en el caso de la mujer se puede presentar una complementación entre el tiempo dedicado al cuidado de los hijos y el tiempo dedicado a trabajar fuera del hogar. Estas razones pueden variar según la clasificación enunciada.

Con el objetivo de contar con una base de datos que contenga la información correspondiente a las variables de interés para la muestra de hogares, se programó una macro en SAS. La misma permite, de una manera automática, seleccionar exclusivamente los hogares de interés y generar un registro para cada pareja conyugal. Una vez seleccionados los hogares, se estudian posibles inconsistencias en algunas variables tales como edad, sexo y escolaridad, revisando la información de otras ondas en los casos que fue posible.

#### IV.- Características de la muestra en estudio

Para realizar el estudio empírico se seleccionaron los aglomerados urbanos: Capital Federal, partidos del Conurbano Bonaerense, Gran Córdoba y Gran Mendoza, debido a que en el primero se presenta la mayor tasa de participación femenina en el mercado laboral del país, en contraposición a Mendoza, uno de los aglomerados con menor tasa. Los aglomerados Córdoba y Conurbano presentan comportamientos intermedios para esta tasa (Tabla IV.1).

La muestra en estudio para cada aglomerado está compuesta por las parejas conyugales (sin importar la legalidad de la unión) que viven en un mismo hogar, entrevistadas en la EPH en la primera onda de 1998. Estas muestras no incluyen aquellas parejas donde el varón tiene ingreso cero o es inactivo o tiene 65 o más años de edad. Se eliminan los varones con ingreso cero porque se presentan un número pequeño de casos, lo que apareja problemas en la estimación. Esta eliminación no representa cambios importantes en los resultados, ya que se eliminan pocos casos.

**Tabla IV.1:** Porcentaje de mujeres con participación laboral, ocupadas y con ingreso cero discriminadas por aglomerado

Aglomerado	Número parejas conyugales	Participación laboral	Ocupadas( en relación con participación)	Ingreso cero (*)
Capital Federal	331	61.6	90.0	47.4
Conurbano	1103	42.1	86.0	65.9
Gran Córdoba	341	40.5	87.6	64.2
Gran Mendoza	471	39.1	98.0	64.6

\*El porcentaje de mujeres sin ingreso no es el complemento del porcentaje de participación debido a la existencia de un número pequeño de mujeres que declaran estar ocupadas pero sus ingresos son cero.

De la Tabla IV.1 se desprende fácilmente que los ingreso cero se deben principalmente a que las mujeres son inactivas y no desocupadas.

Los aglomerados en estudio también presentan un comportamiento diferente en cuanto a la asociación entre los ingresos de fuente laboral cuando ambos cónyuges tienen ingresos

positivos (Gráfico IV.1). Mientras en Capital Federal no parece existir relación entre los ingresos de ambos cónyuges, esta relación podría existir en los tres aglomerados restantes. Así mismo, es distinta la distribución de los ingresos de los varones para las mujeres con ingreso cero (Gráfico IV.1); en el caso de Capital Federal la distribución es más uniforme presentándose ingresos de varones en una proporción similar en un amplio intervalo, mientras que por ejemplo, Conurbano presenta un pico pronunciado para los varones de ingreso medio, comportamiento similar en los otros dos aglomerados pero no tan notorio.

La mediana de ingresos de ambos cónyuges y el rango excluidos los ingresos ceros se presentan en la Tabla IV.2 En ella se puede destacar que Capital Federal posee los ingresos medios más altos, mientras que Mendoza tiene los ingresos medios más bajos pero que el rango de los ingresos del varón es el más grande respecto a los otros aglomerados y además es más del doble que el del ingreso de la mujer. Esto se podría asociar con la participación femenina en el mercado de trabajo, ya que Capital Federal posee el valor más alto y Mendoza el más bajo.

**Tabla IV.2:** Mediana y rango del ingreso laboral de ambos cónyuges por aglomerado

Aglomerado	Iflv		Iflm(sin ceros)	
	Mediana	Rango	Mediana	Rango
Capital federal	1050	9910	700	5970
Conurbano	600	9980	400	6990
Gran Córdoba	500	4980	390	4570
Gran Mendoza	500	10310	400	4255

Como una primera aproximación para medir si existe asociación entre del ingreso laboral de la mujer categorizado en positivo o cero respecto de cada una de las siguientes variables: escolaridad (medescom, medida en las categorías sin escolaridad, primaria incompleta, primaria completa, secundaria incompleta, secundaria completa, mayor que secundaria), edad (agrupada en los intervalos: 10-24, 25-34, 35-44, 45-54, >55), número de hijos menores de 6 años (cont6) y número de hijos entre 6 y 18 años (cont18), se realiza una prueba de independencia ( $\chi^2$ ). Los resultados se presentan en la Tabla IV.3. Se puede destacar que en todos los aglomerados el comportamiento de la variable escolaridad difiere según la mujer tenga ingreso cero o no, con la variable edad sucede lo mismo salvo para el aglomerado Gran Córdoba. En cambio, no existe diferencia para el número de hijos salvo en Córdoba para hijos menores de 6 años y Conurbano para hijos entre 6 y 18 años.

**Tabla IV. 3:** Asociación entre mujeres según ingreso cero o no y variables exógenas

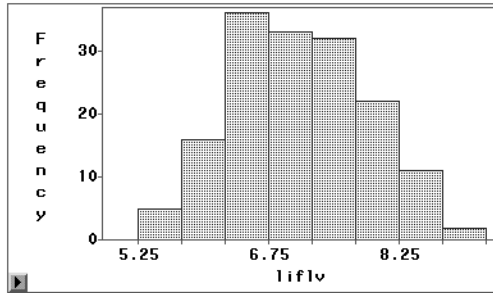
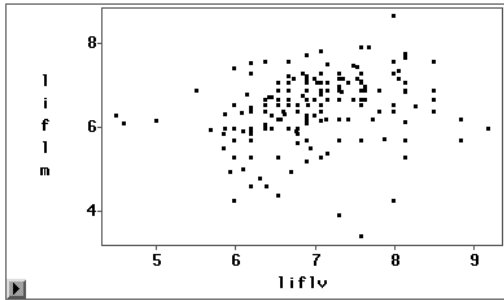
Aglomerado	medescom	Edad	cont6	cont18
Capital Federal	Si	Si	No	No
Conurbano	Si	Si	No	Si*
Gran Córdoba	Si	No	Si*	No
Gran Mendoza	Si	Si	No	No

\*la probabilidad asociada en estos casos se presenta en el limite de rechazo de independencia.

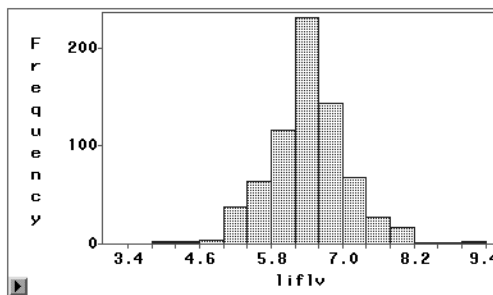
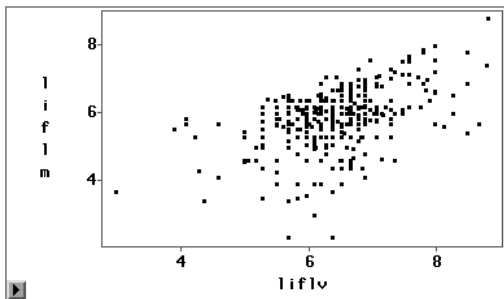


**Gráfico IV.1:** Diagramas de dispersión: liflv vs. liflm (para ambos positivos) e histogramas de liflv dado liflm=0 para los distintos aglomerados.

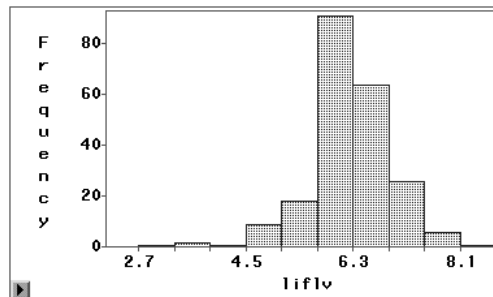
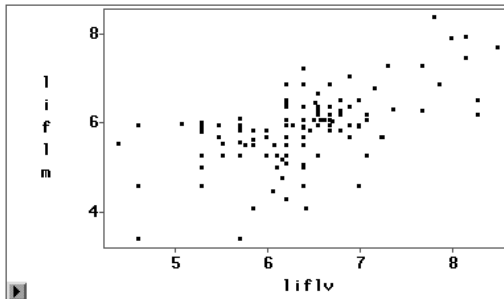
**CAPITAL FEDERAL**



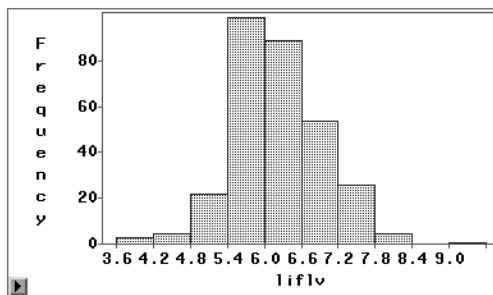
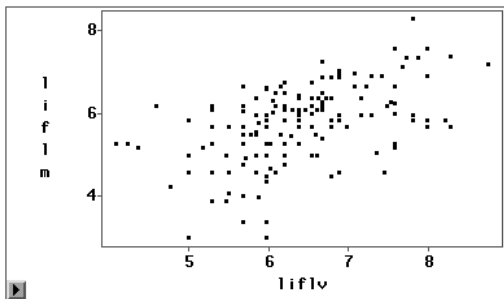
**CONURBANO**



**GRAN CÓRDOBA**



**GRAN MENDOZA**



El breve análisis descriptivo realizado sugiere comportamientos bastante disímiles entre los aglomerados analizados, a esto se podría agregar que pueden existir pautas sociales y culturales distintas entre los mismos. Por estos motivos, es de esperar que la estimación de los modelos pueda diferir entre ellos.

## V. Modelos empíricos

Para modelar el ingreso laboral de la pareja conyugal se considera el modelo (I.9)-(I.10a)-(I.10b), propuesto por Nelson y Olson (1978).

$Y_1$ : vector de la variable endógena para el varón (liflv),

$Y_2$ : vector de la variable endógena para la mujer (liflm),

$X_1$ : matriz de características medibles (variables exógenas) que afectan a  $Y_1$ , (edadv, edad2v, mescov, lifnlv, mescom, lifnlm, cont6, cont18),

$X_2$ : matriz de características medibles (variables exógenas) que afectan a  $Y_2$ , (edadm, edad2m, mescom, lifnlm, mescov, lifnlv, cont6, cont18).

Se presentan las estimaciones realizadas por el método propuesto por Nelson y Olson(1978), también para realizar comparaciones se presentan los resultados de aplicar mínimos cuadrados tres etapas (MC3E) al mismo modelo, considerando sólo las parejas en que ambos cónyuges tienen ingreso positivo (Tabla V.1 ).

Si bien también se realizaron las estimaciones AGLS(Amemiya, 1979) y las estimaciones por el método E-M de la función de verosimilitud normal bivariada truncada, métodos que tienen en cuenta en la segunda etapa de la estimación la covariancia entre los errores, en ambos métodos se presenta el problema que la estimación de dicha covariancia resulta un valor extremadamente grande, lo que invalida todas las estimaciones posteriores. Actualmente se están estudiando las posibles causas de dichos valores. En principio se puede pensar que el alto número de censura (entre 50 al 60 % de ingresos de la mujer es igual a cero), puede ser el origen del problema.

De los resultados encontrados para las estimaciones de los coeficientes del varón teniendo en cuenta la censura se puede distinguir: que su escolaridad es el único coeficiente significativo en todos los aglomerados, que el ingreso de la cónyuge sólo es significativo de signo positivo en el Conurbano, el ingreso de fuente no laboral es significativo de signo negativo en todos los aglomerados salvo en Córdoba y que el coeficiente de la escolaridad de la mujer es positivo y significativo en Córdoba y Mendoza.

En cambio, para el caso de la mujer: el coeficiente del ingreso de fuente laboral, de la escolaridad y del ingreso no laboral de su cónyuge resulta significativo sólo para Capital Federal, con signo negativo, positivo y negativo respectivamente; su escolaridad y el ingreso de fuente no laboral propio son las únicas variables que tienen los coeficientes significativos en todos los aglomerados, de signo positivo y negativo respectivamente, el coeficiente de edad es significativo en Capital Federal, Conurbano y Mendoza; los coeficientes de los hijos menores de seis años son significativos de signo negativo en Conurbano y Córdoba y los hijos entre seis y dieciocho año sólo en el Conurbano.

Estos resultados muestran diferencias respecto a las variables que influyen en la formación según el aglomerado que se analice. Por ejemplo, el ingreso de fuente laboral de la mujer afecta en forma directa el ingreso de fuente laboral de varón sólo en el Conurbano, esto estaría en concordancia con lo enunciado por Lam(1988), en cambio, en los otros tres aglomerados la no influencia de dicha variable avalaría la teoría de acuerdo con la postura de Jacobsen y Rayack(1996).

**Tabla V.1:** Estimación de los coeficientes para ambos cónyuges por aglomerado según distintos métodos.

Variables	Capital Federal		Conurb. bonaerense		Gran Córdoba		Gran Mendoza	
	MC3E	Nel-Ols	MC3E	Nel-Ols	MC3E	Nel-Ols	MC3E	Nel-Ols
<b>VARON</b>								
Intercepto	4.6507*** (0.7752)	4.3294*** (0.0685)	2.5708*** (0.5223)	5.6145*** (0.4968)	2.0647*** (0.7646)	4.1709*** (0.6472)	2.3148** (0.8969)	4.6915*** (0.6845)
Liflm	0.0138 (0.0644)	-0.0087 (0.0440)	0.4058*** (0.0377)	0.0818*** (0.0242)	0.6641*** (0.0880)	-0.0442 (0.0501)	0.4 acuerdo (0.0409)	0.0143 (0.0148)
Edadv	0.0001 (0.0351)	0.0401 (0.0326)	0.0286 (0.0239)	0.0012 (0.0199)	-0.0010 (0.0366)	0.0372 (0.0334)	0.0376 (0.0409)	0.0172 (0.0300)
Edad2v	0.0002 (0.0004)	-0.0003 (0.0004)	-0.0003 (0.0003)	0.0001 (0.0002)	-0.0001 (0.0004)	-0.0004 (0.0004)	-0.0004 (0.0004)	-0.0001 (0.0003)
Medescov	0.0888*** (0.0154)	0.0907*** (0.0132)	0.0700*** (0.0108)	0.0849*** (0.0068)	0.0547** (0.0176)	0.0522*** (0.0107)	0.0621*** (0.0166)	0.0835*** (0.0092)
Lifnlv	-0.0860* (0.0487)	-0.0805* (0.0428)	-0.0081 (0.0360)	-0.0436** (0.0217)	0.0084 (0.0436)	0.0443 (0.0387)	-0.0944* (0.0487)	-1.216*** (0.0311)
medescom	0.0460*** (0.0174)	0.0251 (0.0325)	0.0196* (0.0107)	0.0001 (0.0141)	-0.0152 (0.0210)	0.0779** (0.0356)	0.0072 (0.0168)	0.0242* (0.0139)
Lifnlm	-0.0690 (0.0500)	-0.0074 (0.0397)	-0.0126 (0.0572)	0.0293 (0.0290)	0.1281 (0.1222)	-0.1357 (0.0863)	0.0030 (0.1378)	0.0038 (0.0379)
Cont6	0.0603 (0.0869)	0.0680 (0.1078)	-0.0251 (0.0446)	0.0080 (0.0291)	0.0425 (0.0732)	-0.0904 (0.0643)	-0.0079 (0.0700)	0.0633 (0.0391)
Cont18	0.0740 (0.0608)	0.0235 (0.0456)	-0.0064 (0.0270)	0.0263 (0.0210)	0.0463 (0.0510)	0.0207 (0.0354)	0.0039 (0.0448)	0.0204 (0.0239)
<b>MUJER</b>								
Intercepto	3.8215*** (1.0137)	141.896* (81.940)	0.2179 (0.6789)	44.897 (48.2695)	0.1335 (0.6915)	45.9643 (35.7341)	-1.0317 (1.1676)	-9.7046 (23.5151)
Liflv	0.0211 (0.0916)	-34.867* (18.553)	0.6541*** (0.0604)	-15.801 (11.9160)	0.5625*** (0.0770)	-12.1209 (7.7870)	0.5434*** (0.0802)	-7.3060 (5.4284)
Medescov	0.0096 (0.0197)	3.0061* (1.6677)	-0.0132 (0.0146)	1.135 (0.9263)	-0.0169 (0.01698)	0.6363 (0.4338)	-0.0094 (0.0193)	0.5810 (0.4635)
Lifnlv	0.0372 (0.0550)	-2.631* (1.5836)	-0.0104 (0.0463)	-0.827 (0.6366)	0.0047 (0.0405)	-0.3563 (0.5918)	0.0086 (0.0507)	-0.9031 (0.7278)
Edadm	0.0607 (0.0450)	1.764** (0.7152)	0.0594* (0.0311)	1.682** (0.7679)	0.06497* (0.0337)	0.5553 (0.4202)	0.1226** (0.0537)	1.9691*** (0.4103)
Edad2m	-0.0007 (0.0005)	-0.0136** (0.0046)	-0.0006 (0.0004)	-0.018** (0.0080)	-0.0007 (0.0004)	-0.0067 (0.0050)	-0.0013** (0.0006)	-0.022*** (0.0047)
medescom	0.0983*** (0.0196)	1.433*** (0.4329)	0.0293** (0.0134)	1.197** (0.5189)	0.0782*** (0.0178)	1.2236*** (0.4041)	0.0626*** (0.0178)	0.8349*** (0.2009)
Lifnlm	-0.1019 (0.0627)	-0.846** (0.3304)	-0.0520 (0.0730)	-1.054** (0.4771)	-0.1559 (0.1133)	-2.304*** (0.7641)	0.0413 (0.1530)	-1.594*** (0.5952)
Cont6	-0.290*** (0.1005)	2.348 (2.3428)	-0.0176 (0.0564)	-1.051* (0.5560)	0.0447 (0.0697)	-1.487*** (0.5761)	-0.0111 (0.0793)	0.1632 (0.5151)
Cont18	-0.1665** (0.0694)	0.510 (0.2976)	-0.0562 (0.0345)	-1.111*** (0.4021)	-0.0183 (0.0475)	0.2903 (0.3460)	-0.0507 (0.0489)	-0.4191 (0.2720)

\* significativo al 10%, \*\* significativo al 5%, \*\*\* significativo al 1%.

Los valores entre paréntesis corresponden a los desvíos estándares de los estimadores.

Por otro lado, el ingreso de fuente laboral del varón sólo afecta en forma inversa el de la mujer en Capital Federal.

Otra discusión que se encuentra en la literatura es que si la escolaridad de la mujer afecta los ingresos del varón. Nuevamente los resultados no coinciden entre aglomerados, mientras en Córdoba y Mendoza la escolaridad de la mujer afecta en forma directa el ingreso del varón apoyando la teoría de Becker (1983); en Capital Federal y Conurbano no existe relación. Así mismo, la escolaridad del varón afecta en forma positiva los ingresos de la mujer sólo en Capital Federal.

De los resultados también se desprende que en todos los aglomerados, las estimaciones varían sustancialmente sino se tiene en cuenta la censura de los datos y sólo se trabaja con los ingresos positivos.

## **VI.- Consideraciones finales**

Los resultados empíricos encontrados en este estudio ayudan a comprender principalmente tres aspectos a tener en cuenta al realizar trabajos que traten de explicar la conformación del ingreso individual en la Argentina.

Primero, a través del breve análisis descriptivo realizado en la sección IV se puede apreciar el comportamiento disímil entre aglomerados tanto por el porcentaje de mujeres que no participan en el mercado laboral, como por las características diferente de estas mujeres según sea el aglomerado. Este hecho por supuesto, luego se verá reflejado en la estimación del modelo que intenta explicar la formación del ingreso.

Segundo, especialment respecto de las mujeres si se trabaja sólo con las parejas que tienen ambos ingreso positivo los resultados y por ende las interpretaciones varían fundamentalmente.

Tercero, el alto porcentaje de censura en los datos, entre un cincuenta a sesenta por ciento según sea el aglomerado, hace que se tenga dificultades en la estimación de los parámetros aún cuando se utilicen métodos recomendados tanto en la literatura estadística como econométrica. En este trabajo se presentan los estimadores de Nelson y Olson (1978) porque de los métodos probados es el que mejor funciona, especialmente en cuanto a la variancia de los estimadores, no obstante se debería seguir investigando sobre la real eficiencia de los mismos. Los otros métodos utilizados son aquellos que tienen en cuenta la covariancia entre los errores de las dos ecuaciones, uno es el método AGLS de Amemiya (1979) y otro el método propuesto por Huang et al.(1987) que maximiza la función de verosimilitud de la normal bivariada truncada por medio del algoritmo E-M. En ambos casos la estimación de la covariancia entre los errores de las dos ecuaciones en la segunda etapa se hace extremadamente grande lo que anula todos los pasos posteriores.

Resulta de interés seguir investigando las causas del no funcionamiento de algunos métodos para la estimación del modelo. Por esto se piensa estudiar más profundamente la consistencia del modelo y por otro lado realizar simulaciones cambiando el grado de censura en forma aleatoria para ver si esto modifica los resultados.

## **Referencias Bibliográficas**

- Akerlof, G (1997) "Social distance and social decisions" *Econometrica*, 65, 1005-1027.
- Amemiya, T. (1974) "Multivariate regression and simultaneous equation models with dependent variables are truncated normal". *Econometrica*, 42, 999-1012.
- , T. (1978) "The estimation of a simultaneous-equation generalized probit model". *Econometrica*, 46, 1193-1205.

- , T. (1979) "The estimation of a simultaneous-equation Tobit model" *International Economic Review*, 20, nº1, 169-181.
- Becker, Gary (1973) "A Theory of Marriage:Part 1", *Journal of Political Economy*, Vol.81, Nº 4.
- (1983): *El Capital Humano*, Alianza Universidad, Madrid.
- Blaconá, M.T., Garcia, M.del C., Borgognone, M.G., Bussi, J. y Pellegrini, J.L. (2001)" Consideraciones metodológicas sobre la estimación econométrica de las ecuaciones de ingresos de los integrantes de la pareja conyugal".Reunión Anual de la Asociación Argentina de Economía Política. CD-ROM pág. 31
- Boulier, Bryan L. and Rosenzweig, Mark (1984): "Schooling, Search, and Spouse Selection: Testing Economic Theories of Marriage and Household Behavior", *Journal of Political Economy*, Vol.92m nº 4.
- Dempster,A., Laird, N. y Rubin, D. (1977) "Maximun Likelihood from incomplete data via the EM algorithm". *Journal of the Royal Statistical Society*, B, 39, 1-22.
- Griliches, Z., (1977): "Estimating the Returns to Schooling: Some Econometric Problems", *Econometrica*, 45(1), pp 1-22.
- Heckman, J. (1974) "Shadow price, market wages and labor supply". *Econometrica*, 42, 679-694.
- Huang, C., Sloan, F. y Adamache, K. (1987) "Estimation of seemingly unrelated Tobit regressions via EM algorithm". *Journal of Business & Economic Statistics*, 5, nº 3, 425-430.
- Jacobsen, J. y Rayack, W. (1996): "Do Men Whose Wives Work Really Earn Less?", *American Economic Review*, Vol 86(2), May, pp 269-273.
- Lam, D. (1988) "Marriage Markets and Assortative Mating with Household Public Goods", Theoretical Results and Empirical Implications, *The Journal of Human Resources*, XXIII. 4, pp. 462-487.
- Lee, L. F. (1993) "Multivariate Tobit models in econometrics". *Estimation of limited dependent variable models under rational expectations*. Maddala, S. Rao, C. R. y Vinod, H. Elsevier Science Publishers.
- Mincer, J. (1970) "The distribution of labor income: a survey; With special reference to human capital approach", *Journal of Economic Literature*, Vol. VIII, Nº1.
- Nelson, F. y Olson, L. (1978) "Specification and estimation of a simultaneous-equation model with limited dependent variables". *International Economic Review*, 19,nº3, 695-709.
- Tobin, J. (1958): " Estimation of Relationships for Limited Dependent Variables", *Econometrica*, 26, pp. 24-36.