



Universidad de la República
Facultad de Ciencias Sociales
DEPARTAMENTO DE ECONOMIA

Documentos de trabajo

El sesgo de selección en la actividad de jóvenes y mujeres

Rafael Diez de Medina

Documento No. 12/92
Diciembre, 1992

EL SESGO DE SELECCION EN LA ACTIVIDAD DE JOVENES Y MUJERES (*)

Rafael Diez de Medina (**)

BANCO CENTRAL DEL URUGUAY
BIBLIOTECA

RESUMEN

El mercado de trabajo de Uruguay ha presenciado un crecimiento acelerado de las tasas de actividad de sus grupos secundarios- mujeres y jóvenes- en las últimas dos décadas. Este artículo parte de esta realidad ya consolidada y estima los determinantes que llevan a ambos grupos a entrar en el mercado, enfatizando en los problemas metodológicos de estimación que acarrea el ignorar el llamado *sesgo de selección* a la hora de estudiar este tema. A través de la modelización de jóvenes, por un lado, y mujeres por otro, se ensayan dos maneras de aislar dicho sesgo. Se identifican los determinantes de los salarios, en el caso de los jóvenes y los de la diferente dedicación laboral, en el de las mujeres utilizando para ambos modelos simultáneos.

ABSTRACT

SELECTIVITY BIAS IN YOUTH AND FEMALE PARTICIPATION

The Uruguayan labor market has witnessed an accelerated growth in the participation rates of its secondary labor force namely, youth and female, in the last two decades. This paper assumes that this reality is consolidated and estimates the determinants that lead both groups to participate, stressing on the methodological issues that arise if the so-called *sample selection bias* is ignored when tackling this problem. By modelling the youth and female behavior, two ways of handling this problem are submitted. The determinants of wages, in the case of youth, and of the different hourly dedication in female labor force, are identified by using simultaneous models.

(*) Este trabajo fue realizado en el marco del área de Econometría del Departamento de Economía de la Facultad de Ciencias Sociales.

(**) Docente del Departamento de Economía (Facultad de Ciencias Sociales) y de la Facultad de Ciencias Económicas, Universidad de la República e investigador de la Oficina de CEPAL en Montevideo.

Introducción.

Entre las principales características del mercado de trabajo urbano del Uruguay de las últimas dos décadas, se destaca especialmente la incorporación creciente de los grupos generalmente denominados "secundarios", es decir, mujeres y jóvenes, a la fuerza laboral. Dichos grupos muestran tasas de crecimiento en su participación que podrían calificarse de espectaculares teniendo en cuenta otras realidades y la propia trayectoria temporal de estas variables en el país.

Estudiar con atención este fenómeno parece tener relevancia a la hora de explicar el comportamiento del mercado en lo que hace a las consecuencias que, sobre él, tiene esta masiva incorporación. Así, se puede intuir la importancia que sobre la institucionalidad del mercado laboral tiene el hecho de que crecientes grupos femeninos compitan cada vez más con los hombres, poniéndose de manifiesto un importante fenómeno de discriminación y segregación por sexo (Rossi y Diez de Medina, 1989; Rossi y Bucheli, 1987), a la vez que una creciente demanda por mecanismos de inserción para los grupos más jóvenes que se ven muchas veces desplazados a causa de asimetrías entre oferta y demanda y por factores de rigidez como cierta legislación sobreprotectora no adaptada a las nuevas circunstancias. El estudio de la actividad y sus determinantes se impone también si se considera que la creciente actividad viene acompañada de altas tasas de desempleo abierto y de subutilización en la mano de obra, particularmente en estos grupos especiales.

Se han ensayado varias explicaciones del fenómeno que apuntan a la hipótesis del *added worker*, es decir que ciertos mecanismos desencadenantes impulsaron a los hogares a recomponer caídas en sus ingresos y, a través de una decisión de asignación interna de sus recursos humanos, se habría optado por incorporar grupos que, de haber mediado otras circunstancias, no hubieran salido al mercado. Sin embargo, la persistencia de elevadas tasas de actividad aseguran que la explicación deberá ser realizada en otros términos, puesto que, desde ya, el fenómeno se ha constituido en estructural en la realidad uruguaya.

Si bien el signo de la tendencia hacia la incorporación creciente de la mujer al mercado no llama la atención, puesto que se presenta como una característica en casi todos los mercados con cierto desarrollo relativo, lo que sí se presenta como un desafío es la explicación a la aceleración con la que se da en Uruguay. Mientras en 1969, la tasa promedio anual de actividad de la mujer era del 27%, en 1991 la misma ascendió a casi el 47%. Por otra parte, en lo que hace a la participación de los grupos jóvenes (entre 14 y 24 años), mientras en 1973 el porcentaje de activos en el total de la población de esta edad era de poco más del 40%, en 1991 se alcanza a más del 55%, con mayor aceleración en los grupos de entre 14 y 19 años, cuya tasa ha aumentado la cuarta parte en el mismo lapso.

El objetivo principal de este artículo es el de presentar un análisis de corte transversal que intenta modelizar el comportamiento microeconómico de ambos grupos para la ciudad de Montevideo. Esta metodología se aplica sobre datos correspondientes al segundo semestre de 1991, con el fin de explicar las decisiones que se revelan en el mercado laboral.¹ A tales efectos, se estudia con especial énfasis el problema del llamado "sesgo de selección" (Roy, 1951; Heckman, 1979) que surge en toda estimación que se basa en subpoblaciones de interés que, aunque muchas veces se ignora, no se extraen aleatoriamente a la hora de analizarlas. En efecto, los datos se refieren a los jóvenes y mujeres activos, y son ellos (y no el encuestador) quienes toman la decisión de participar en el mercado de trabajo. Si no se controla por la posibilidad de sesgo de selección, y no se ajustan los métodos empíricos en consecuencia, existe el riesgo de inconsistencia en las estimaciones.

En la primera sección, el artículo presenta los distintos instrumentos microeconómicos utilizados, realizando una breve revisión de la literatura más reciente que incorpora el problema de la autoselección. En el caso de la modelización de la actividad femenina, se opta por la especificación de un modelo *probit* bivariado con el fin de incorporar, además de la decisión de participar, la de la diferente dedicación horaria. En el caso de la actividad de los grupos jóvenes, la modelización sigue la línea de otros trabajos (Diez de Medina, 1992; Diez de Medina, Rossi, 1991) utilizando un sistema simultáneo que incluye la estimación de una ecuación salarial a los efectos de probar la hipótesis de existencia de sesgo de selección.

En las secciones segunda y tercera, se analizan los resultados a los que se arribó, luego de la aplicación de la metodología presentada en la primera parte, utilizándose para su estimación, los microdatos muestrales de la Encuesta Continua de Hogares que realizó la Dirección General de Estadística y Censos en el segundo semestre de 1991 para el departamento de Montevideo.

1. El sesgo de selección y la actividad

El sesgo de selección o autoselección surge cuando la extracción de una subpoblación cualquiera de interés no se realiza con equiprobabilidad para todos sus elementos, ya sea por la forma de extracción estadística de la muestra, como por las propias decisiones de los agentes estudiados o ambas causas a la vez. En el muestreo aleatorio se cuenta siempre con la información de la probabilidad *a priori* de extracción de cada integrante de la

¹ En dicho período ya se juzgan consolidadas las altas tasas de actividad mencionadas anteriormente.

muestra y por lo tanto, de existir igual probabilidad, una mayor muestra brindará mejores estimaciones. Sin embargo, de no conocerse esa "regla" por la cual se extraen aleatoriamente sus integrantes, el análisis puede describir características de la muestra pero no de la población, aún cuando aquélla sea muy grande: se está en presencia de un sesgo de selección. Los ejemplos abundan: la observación de salarios u horas trabajadas sólo se tiene para aquellas mujeres o jóvenes que efectivamente han optado por participar, puesto que el salario que perciben es mayor que el de reserva, y lo hacen exitosamente; el análisis de los desempleados sólo se hace con aquellas personas que mostraron efectivamente estar cesantes, etc. Como se observa, todos esos ejemplos involucran un proceso previo de decisión por parte de los individuos observados que los hizo incluirse o no en el análisis (de ahí el nombre de autoselección).

La existencia de este sesgo en las estimaciones econométricas es señalado ya por Roy (1950) en su célebre artículo sobre la elección por los agentes de diferentes ocupaciones dentro de un conjunto de oportunidades. Allí, se explicita que la decisión por una ocupación u otra se ve estadísticamente a través de las personas que efectivamente eligieron y, por su parte, no se observan las habilidades "latentes" que los hacen más aptos para cada actividad. De lo que se trata entonces es de aprovechar la relación que existe entre la variable latente no observada y la efectivamente visible. Fue Heckman (1974, 1979, 1987) quien formalizó la identificación del sesgo y propuso una metodología para su corrección.

Para explicitar el sesgo del punto de vista econométrico, se puede plantear una muestra aleatoria de N individuos, donde para cada uno de ellos se plantea el sistema:

$$Y_{1i} = X_{1i} \cdot \beta_1 + V_{1i} \quad (1)$$

$$Y_{2i} = X_{2i} \cdot \beta_2 + V_{2i} \quad \text{con: } i = 1, 2, \dots, N \quad (2)$$

$$E(V_{1i} \cdot V_{2i}) = \sigma_{12} \quad , \quad E(V_{1i}^2) = \sigma_{11} \quad , \quad E(V_{2i}^2) = \sigma_{22} \quad , \quad E(V_{ij}) = 0$$

$$\text{y: } E(V_{im} \cdot V_{jn}) = 0 \quad \text{para } m \neq n$$

En este modelo, Y_{ij} es, por ejemplo, el ingreso salarial del individuo i -ésimo que se observa únicamente si éste es un ocupado, es decir si Y_{2i} (diferencia entre el salario ofrecido en el mercado

y su salario de reserva²) es positivo. Por su parte, X_{ji} es un vector de variables explicativas y β_i es el vector asociado de parámetros. Así, los años de educación, el sexo y la experiencia serían ejemplos de variables incluidas en el vector X_{1i} , pues determinan los salarios logrados. Por otro lado, la cantidad de niños en el hogar, el estado civil y también los años de educación, serían ejemplos de variables del vector X_{2i} , pues influirían en la decisión de participar en el mercado o no.

Se desea estimar la ecuación (1), por ejemplo, una ecuación salarial, pero hay algunos individuos para los cuales no se observa dicha variable.

Mientras que la expresión de la regresión poblacional es:

$$E(Y_{1i}/X_{1i}) = X_{1i} \cdot \beta_1$$

en la muestra, la expresión va a depender de la forma en que fueron elegidos los elementos a investigar. Supóngase que sólo se observan salarios para las mujeres que optaron por participar, es decir, cuando Y_{2i} , la diferencia entre el salario ofrecido y el de reserva fuera positivo. Entonces, la esperanza condicional de la muestra efectiva tendrá en cuenta la regla de selección:

$$E(V_{1i}/X_{1i}, Y_{2i} \geq 0) = E(V_{1i}/V_{2i} \geq -X_{2i} \cdot \beta_2) \quad (3)$$

por lo que, sustituyendo (3) en (1) y tomando esperanzas condicionales a la regla de selección se obtiene (4):

$$E(Y_{1i}/X_{1i}, Y_{2i} \geq 0) = X_{1i} \cdot \beta_1 + E(V_{1i}/V_{2i} \geq -X_{2i} \cdot \beta_2) \quad (4)$$

donde el segundo término incluye explícitamente el sesgo de selección que lleva a que, de ignorarlo, se cometan errores de omisión de variables. Si (V_{1i}, V_{2i}) se suponen con distribución normal bivariada, el término (3) tiene la forma de una esperanza de una distribución normal truncada:

$$E(V_{1i}/V_{2i} \geq -X_{2i} \cdot \beta_2) = -\sigma_{12}/\sqrt{\sigma_{22}} \cdot \lambda_i \quad (5)$$

siendo: λ_i (λ) el inverso del ratio de Mill, es decir:

$$\phi(Z_i)/\Phi(Z_i)$$

² El concepto de salario de reserva (W^*) es un concepto de costo de oportunidad del tiempo libre de un individuo: el umbral mínimo por el cual éste aceptaría ofrecerse en el mercado. Obviamente, si el salario de mercado es mayor a este umbral, el individuo participa, mientras que si es menor, el costo de oportunidad es considerado alto y se opta por el tiempo libre.

donde:

$$Z_i = -X_{2i} \cdot \beta_2 / \sqrt{\sigma_{22}} \quad y:$$

$\phi(\cdot)$ y $\Phi(\cdot)$ representan respectivamente, las funciones de densidad y distribución de una normal estandarizada.

Heckman (1979) propone un método bietápico de identificación y especificación del sesgo. En una primera etapa se estiman los parámetros de la probabilidad de la regla de selección (en el ejemplo, que $Y_{2i} \geq 0$) mediante un modelo *probit* y se computa para cada observación la variable λ . En una segunda etapa, se introduce λ explícitamente como variable explicativa de la ecuación (1) y se realiza la prueba de hipótesis sobre el coeficiente asociado a dicha variable. Si el coeficiente resulta significativamente distinto de cero, es que efectivamente hay un sesgo de selección.

La explicitación de este problema se realiza basándose generalmente en modelos de elección discreta (ya sea binario o multidimensional), puesto que la identificación del sesgo parte de estimar la probabilidad de que efectivamente se cumpla la regla de selección y se combina con ecuaciones lineales que generalmente corrigen explícitamente el sesgo de selección estimadas por mínimos cuadrados ordinarios. Este método es el utilizado para la estimación del modelo explicativo de la actividad de los jóvenes en Montevideo.

Sin embargo, para la explicación de la actividad femenina que se presenta se opta por un modelo *probit* bivariado, del tipo del utilizado por Greene (1984) y Kiefer (1982). Este modelo simultáneo tiene la ventaja de incorporar el sesgo de selección de una manera más completa, estimándose mediante máxima verosimilitud con información completa (FIML). La formulación se adecua especialmente al caso de la participación femenina, puesto que implica no sólo modelizar la probabilidad de actividad, sino, simultáneamente, estimar la probabilidad de que participe en forma de tiempo completo o parcial. Tomando en cuenta que la mujer posee la particularidad de actuar en el mercado en forma parcial en un número importante de casos, a causa principalmente de la compatibilización con roles de cuidado de los niños y tareas domésticas, se juzga interesante optar metodológicamente por un instrumento que incluya explícitamente dichos elementos.

El modelo planteado será:

$$Y_{1i}^* = X_{1i} \cdot \beta_1 + V_{1i} \quad (6)$$

$$Y_{2i}^* = X_{2i} \cdot \beta_2 + V_{2i} \quad (7)$$

con:

$$Y_{ji} = 1 \text{ si } Y_{ji}^* \geq 0 \quad \text{y: } Y_{ji} = 0 \text{ en otro caso, para: } j = 1, 2$$

$$E(V_{1i}) = E(V_{2i}) = 0, \quad \text{Var}(V_{1i}) = \text{Var}(V_{2i}) = 1, \quad \text{cov}(V_{1i}, V_{2i}) = \rho$$

siendo Y_{ji}^* variables latentes no observadas y Y_{ji} variables observadas con la única restricción que (Y_{1i}, X_{1i}) se observa sólo cuando $Y_{2i}=1$ y ρ (rho), la correlación entre los residuos de (6) y (7).

En nuestro caso, la elección entre tiempo completo y tiempo parcial en la dedicación femenina (Y_1) se da únicamente cuando la mujer ha decidido participar en el mercado laboral ($Y_2=1$), por lo que los datos observados no son una muestra aleatoria de todos los posibles casos. Se deseará estimar conjuntamente ambas probabilidades.

¿Dónde está explicitado el sesgo de selección? Al no contarse con ecuaciones lineales no se puede utilizar el método bietápico de Heckman y por lo tanto no se hará uso del mismo. La selección dependerá de la significatividad de la estimación de rho. Si éste fuera cero, se estaría en presencia de dos probabilidades independientes, por lo que no existiría sesgo de selección en la intensidad de la participación de la mujer. En cambio, de ser significativamente distinto de cero, la selección estará presente. La ventaja estará en la utilización de un método consistente como el de FIML y las pruebas de significación utilizadas serán: el estadístico de Wald y el test de la razón de verosimilitud (Engle, 1984), en virtud de que el LM (multiplicador de Lagrange) es de muy difícil cálculo (Kiefer, 1982). Cabe señalar, finalmente, que el estimador utilizado para incorporar al modelo probit bivariado el problema de la selección ha sido el propuesto por Wynand y Bernard (1981) y el algoritmo utilizado por FIML, el de Davidson-Fletcher-Powell (DFP).

2. La participación de los jóvenes.

2.1 El modelo planteado

Para realizar la modelización de la participación de los jóvenes se optó por seguir el método bietápico de Heckman. En primer lugar se estimó un modelo explicativo del tipo *probit* (Amemiya, 1981 ; Maddala, 1983 ; Gourieroux, 1989 ; Killingsworth, 1983). para la probabilidad de actividad, en donde se identificaron distintos grupos de variables independientes.

A los efectos de explicar los condicionantes de la participación o no en el mercado laboral por parte de los grupos menores de 29 años, se optó por discriminar dos subgrupos: aquéllos entre 14 y 24 años y los que tienen entre 25 y 29 años. ¿Por qué esta distinción? Principalmente por dos razones: el hecho de que el propio ciclo vital y las características de autonomía o no del hogar de origen diferencian los grupos más jóvenes del resto y, por otro lado, una razón práctica: las personas de entre 14 y 19 años son numéricamente pocas y por tanto conducen a estimaciones con gran variabilidad. Es de señalar, asimismo, que los jóvenes del segundo grupo de edad comparten características comunes a los adultos aunque, dada la estructura etaria del país, son considerados dentro de la categoría de juventud.

El modelo que se plantea es el siguiente:

$$W_i = f(X_{1i}, \beta_1, \lambda_i, V_{1i}) \quad (8)$$

$$P(Y_{2i}=1) = g(X_{2i}, \beta_2 + V_{2i}) \quad (9)$$

donde Y_2 es una variable *dummy* que adopta el valor 1 cuando el individuo i -ésimo es un activo y 0 cuando es inactivo y X_2 es la matriz de variables que resultaron adecuadas para la explicación de la actividad. Estas últimas son de dos grupos: características personales y del hogar, siendo éstas de gran importancia para la explicación de la decisión de participación de los jóvenes, poniendo de manifiesto que la opción por dicha actividad se da en un proceso intrafamiliar de maximización de utilidad. En la segunda etapa, se estima una clásica ecuación salarial minceriana (Mincer, 1974) con el fin de probar la existencia de sesgo de selección a través de la significación de la variable λ que se computa a partir de la ecuación determinante de la probabilidad de actividad.³

³ Al respecto:

$$E(W_i/W_i > W^*) = X_{1i} \cdot \beta_1 - [\sigma_{12}/\sqrt{\sigma_{22}}] \cdot \lambda_i$$

donde σ_{12} es la covarianza entre los residuos de ambas ecuaciones. Es interesante observar el significado de todo el coeficiente

2.2 Los resultados.

El modelo se estimó a partir de una formulación *probit* aplicado a una muestra de 3688 jóvenes menores de 29 años, de los cuales, 1576 eran ocupados y sus resultados se mostraron significativos aplicando el *test* del ratio de verosimilitud. En lo que refiere a la estimación de la segunda etapa (ecuación salarial) se aplicaron mínimos cuadrados ordinarios a la submuestra de los ocupados.

En primer término se estimaron las probabilidades de actividad en los dos grupos planteados: 14 a 24 años, por un lado y 25 a 29, por otro, cuyos resultados se presentan en el Cuadro 1.

Las variables seleccionadas para explicarlas fueron:

1) las características personales de los jóvenes: la edad, el sexo, el estado civil, la educación (en años), la asistencia o no a establecimientos de educación y el hecho de ser o no un jefe de hogar.

2) las determinantes que provienen de los hogares en los que vive el individuo: el ingreso per cápita por adulto equivalente (se contó a cada menor de 14 años como medio adulto), las tasas de desempleo, actividad y dependencia en el hogar y el número de perceptores de ingresos dentro del hogar. La tasa de dependencia del hogar se computó como la cantidad de inactivos por ocupado dentro del hogar y busca captar otro factor de presión sobre la decisión hogareña por la participación de sus integrantes.

La significatividad de cada variable es analizada a partir del estadístico de Wald (WALD), el que es válido en modelos de tipo no lineales estimados por máxima verosimilitud como el presentado (Engle, 1984) y se define como:

$$WALD = \beta_j^2 / \hat{V}ar(\beta_j) \quad \text{con: } WALD \sim \chi^2_1 \quad \text{bajo la } H_0) \beta_j = 0$$

Como se desprende del Cuadro 1, las fuerzas inhibitoras de la actividad en el caso de los más jóvenes son la asistencia a la educación y el ingreso del hogar ya que, a mayor cuantía de éste, menor la probabilidad de que el joven menor de 24 se ofrezca en el mercado. Esta última variable no aportó explicación en el caso de los de mayor edad, debido a las características de su ciclo vital y proceso de autonomía llevan a que el ingreso del hogar no sea tan significativo en los determinantes de su participación o no. La variable que señala el carácter de jefe o no del hogar cobra mayor

asociado a lambda, ya que de ser positivo, estaría indicando que los individuos con alta propensión a tener mejores salarios, también tendrían un costo de oportunidad alto para participar (alto salario de reserva).

importancia en el segundo grupo, puesto que es precisamente en él que se produce el proceso de autonomización y la cantidad de jóvenes que asumen carácter de jefe es mayor y más determinante en su decisión de actividad.

Son importantes los factores asociados a la situación económica del hogar al que pertenece el joven. Así, en la medida de que existe una mayor tasa de actividad en el hogar, entendiéndose por tal al porcentaje de activos en el total de mayores de 14 años, existe mayor actividad. Esto se une, en el mismo sentido, a una presión que opera la tasa de desempleo hogareña y la tasa de dependencia ⁴ (aunque esta última parece no ser significativa en el grupo de menor edad).

Es interesante notar el comportamiento de los años de educación. Mientras en el primer grupo no resultó tener un impacto importante en la participación, en el segundo alcanza una mayor relevancia. Esto es compatible con el hecho de que es en este grupo etario donde varios niveles de enseñanza llegan a su término y por lo tanto el joven que finaliza se ofrece (existosamente o no) al mercado laboral.

De los modelos expuestos, se puede concluir que la situación económica del hogar es muy relevante en la participación de los jóvenes menores de 24 años, por lo que se reafirma allí la vigencia de la teoría del *added worker*, o del "trabajador añadido" por la cual las condiciones de bajos ingresos hacen que el hogar decida por una mayor participación de sus integrantes "secundarios". Esto no ocurre en el caso de los jóvenes mayores de 25 años.

En segundo lugar, se estimaron las ecuaciones salariales por mínimos cuadrados ordinarios con el fin principal de estimar el sesgo de selección. Como se observa del Cuadro 2 la autoselección no parece ser significativa (véase coeficiente asociado a λ), especialmente en los mayores de 24 años, lo que es compatible con las altas tasas de actividad que imperan en estos grupos. En consecuencia, la no aleatoriedad de su extracción parece no ser determinante para invalidar, por ejemplo, un análisis salarial. En los de menor edad, la selección sí alcanza cierta significatividad.

⁴ Es de señalar que muchas veces el coeficiente asociado a esta variable tiene signo negativo, especialmente cuando se ubica en hogares de bajos ingresos donde la cantidad de niños inhibe la actividad de sus miembros a raíz de los costos de oportunidad más altos que se dan.

Cuadro 1

DETERMINANTES DE ACTIVIDAD JUVENIL
Montevideo. Segundo semestre de 1991.

Modelo PROBIT

Variable dependiente: 1= Activo, 0= Inactivo.

	Coeficientes	
<u>Variables explicativas</u>	<u>14 a 24 años</u>	<u>25 a 29 años</u>
Constante	-7.12677 (209.9)	-7.76513 (17.9)
Edad	0.22934 (184.1)	0.05997 (1.1)
Sexo (1=hombre,0=mujer)	0.64119 (69.6)	0.98433 (19.36)
Asistencia a educación (1=asiste,0=no asiste)	-1.2945 (179.6)	-0.8626 (9.8)
Jefatura (1=jefe, 0=no jefe)	1.7998 (24.7)	2.35838 (34.2)
Años de educación	0.0591 (11.2)	0.07517 (5.2)
Estado Civil (1=soltero,0=no soltero)	0.97731 (57.8)	1.06695 (28.1)
Número de perceptores	0.19137 (28.4)	0.50804 (36.4)
Ingreso del hogar (en logaritmos)	-0.2177 (11.6)	0.09298 (0.4)
Tasa desempleo en el hogar	0.51775 (6.7)	0.79359 (3.1)
Tasa de dependencia en el hogar	0.01692 (0.4)	0.38741 (6.8)
Tasa de actividad en el hogar	4.18485 (225.0)	5.73106 (96.1)
Numero observaciones:	2710	978
Log.de verosimilitud:	-737.62	-150.14
Porcentaje predicho:	89%	94%

(.)= Estadístico de Wald distribuido chi-cuadrado.

Cuadro 2

ECUACIONES SALARIALES CON SESGO DE SELECCION

Montevideo. Segundo semestre de 1991

Variable dependiente: log.del ingreso salarial

Método: OLS

Coeficientes

<u>Variables explicativas</u>	<u>14 a 24 años</u>	<u>25 a 29 años</u>
Constante	0.04453 (-0.177)	-0.6373 (-1.49)
Años de educación	0.11131 (13.824)	0.09384 (6.587)
Horas trabajadas(log)	0.64911 (17.911)	0.57038 (11.731)
Experiencia	0.10566 (5.5186)	0.13750 (4.9902)
Exp ²	-0.0006 (-0.420)	-0.0033 (-3.040)
Sexo	0.14047 (4.1294)	0.24482 (6.0519)
Ingreso del hogar (en log)	0.24698 (9.3586)	0.35258 (10.885)
LAMBDA	0.08163 (1.7102)	-0.1274 (-1.226)
Número de observaciones	946	630
R ² corregido:	0.446	0.420

(.)= estadístico t

3. La intensidad de la actividad femenina

3.1 El modelo planteado

Como ya se señalo en la primera sección, se optó por realizar una estimación simultánea que tomara en cuenta conjuntamente la decisión de participar ($Y_2=1$) y, condicional a la actividad, la distinta dedicación, dicotomizando la decisión en: participar por lo menos 30 horas semanales ($Y_1=1$) y menos de 30 horas ($Y_1=0$)⁵.

Con respecto a la decisión de participar, al igual que en el caso de los jóvenes, se eligieron dos clases de variables explicativas: las que hacen a las características personales, y aquéllas que son representativas de la situación del hogar.

En el caso de la mujer, se admite que cuanto mayor sea el número de niños en el hogar, mayor será el salario de reserva y menor la probabilidad de ser activa. Para probar este aserto, se tomaron tres variables: número de niños de entre 0 y 5 años, número de niños de entre 6 y 10 años y número de niños de entre 11 y 14 años de edad.

Por otra parte, el número de desempleados en el hogar propicia la actividad de la fuerza de trabajo secundaria del hogar. Se incluyó la variable *dummy* para la presencia o no de servicio doméstico con la intención de captar los efectos de contar con ayuda en los quehaceres domésticos que facilitan la actividad de la mujer, especialmente en aquéllas que provienen de hogares con ingresos medios y altos.

Asimismo, se utilizaron cuatro variables *dummies* que intentan recoger el efecto del ciclo vital dentro de la etapa laboral de la mujer: cuando la mujer tiene entre 14 y 19 años, entre 20 y 29 años, entre 30 y 39 años y finalmente, entre 40 y 49 años, siendo la omitida el grupo de 50 y más. Por otra parte, se construyó la variable "años de educación"⁶. Esta se juzga muy importante como propiciadora de la actividad femenina, puesto que una mayor incorporación de capital humano, eleva los salarios ofrecidos y por

⁵ La variable Y_2 de acuerdo a (7) es una variable dicotómica que toma el valor unitario cuando se opta por participar y es nula en el resto de los casos. Con respecto a la decisión vale realizar la precisión que participar no es sinónimo de tener éxito en la inserción, sino que se incluyen en este concepto el importante número de mujeres que buscan empleo.

⁶ Al respecto se realizó una recodificación de las variables "nivel" y "años aprobados dentro de cada nivel" que recoge la encuesta para traducirlas en términos de años totales aprobados de estudio.

Por otra parte, la construcción del test de la razón de verosimilitud (LR) reafirma la conclusión a la que conduce el de Wald. La hipótesis nula de inexistencia de sesgo de selección se probará realizando el LR con:

$L(\hat{\omega}) =$ suma de las funciones de verosimilitud por separado y;
 $L(\hat{\Omega}) =$ verosimilitud del modelo probit bivariado.

La función de verosimilitud que incorpora la restricción de H_0 , hipótesis nula, será igual a la suma de los logaritmos de las funciones correspondientes a los modelos probit que surgen para cada decisión en particular. La función $L(\hat{\Omega})$ será la verosimilitud cuando no existe a priori el supuesto de que las decisiones son independientes y por tanto no supone que ρ sea nulo, con lo que se está frente al modelo probit bivariado. Como $-2\ln(LR)$ tiene una distribución chi-cuadrado con un grado de libertad, la ausencia de correlación entre las decisiones podrá ser fácilmente probada. En este caso particular, la suma de los logaritmos de las funciones de verosimilitud por separado alcanzó a: -5575, por lo que el test arroja un valor de 6.⁷ Este resultado apoya fuertemente la existencia de sesgo de selección.⁸

3.3 Comentarios

Los signos de las relaciones son los esperados a priori, donde los factores personales como los años de educación, el estado civil y el carácter de ser jefe influyen en la decisión de actividad. Por otra parte, las características del hogar, especialmente en lo referente a la existencia de niños y de ayudas en el hogar (servicio doméstico), así como los determinantes económicos (desempleo, perceptores) influyen también en la forma esperada.

Lo interesante de la estimación simultánea está en el hecho de que los efectos de varias variables, como las de la existencia de niños y algunas otras características del hogar, son determinantes en la decisión de ajustar el horario de trabajo y no tanto en la decisión de participar.

Del modelo estimado presentado en el Cuadro 3 se pueden extraer varias conclusiones. En primer lugar, la decisión de

⁷ Así es como $-2\ln(LR) = -2[(-3486.3 - 2088.7) + 5572] = 6$ cae dentro de la región crítica de una distribución chi-cuadrado con un grado de libertad. Como se observa, el test reafirma lo ya señalado por el de WALD en cuanto a la existencia de sesgo de selección en la participación de la mujer.

⁸ Véase en Diez de Medina, Rossi (1991) donde se realiza una estimación alternativa basada en el método de Heckman utilizado en este trabajo para los jóvenes y se acepta también la existencia de sesgo de selección.

participar en la fuerza de trabajo por parte de la mujer se ve influida lógicamente por la edad. En los grupos menores, donde no hay una gran acumulación de capital humano ni experiencia es lógico suponer que el salario de mercado sea bajo y por tanto se tendrán pocos incentivos a la actividad. El factor de asistencia a centros de educación se presenta como una alternativa inhibidora. Sin embargo, al aumentar la edad el signo de la relación es positivo. Cabe señalar que en otras realidades, se da un cambio de signo en el grupo asociado a la edad de procreación pues se supone una interrupción generalizada en la vida activa. No parece ser este el caso en la realidad de Montevideo, donde la existencia de niños opera negativamente a la actividad pero, especialmente, influyen en la dedicación e intensidad y no tanto en la actividad misma. Este efecto se traduce en que, por ejemplo, la existencia de niños no opera tanto sobre la actividad misma, sino más bien sobre la dedicación horaria (véase por ejemplo el coeficiente asociado a los niños menores de 5 años). Esto se compatibiliza con la legislación que facilita la actividad parcial en las mujeres que recién han tenido un niño.

Por otra parte, en la decisión de dedicarse más tiempo influyen otras variables como el hecho de percibir, además, otra fuente de ingreso. En el caso de las mujeres que también son jubiladas, se aprecia una tendencia a una inserción menor en términos de horario, mientras que la percepción de pensiones no es determinante a causa obviamente de su reducido valor. El carácter de asalariada privada o de patrón de las activas influyen positivamente hacia una inserción en forma de tiempo completo, mientras que el compartir la actividad con el estudio lo hace en forma negativa.

4. Conclusión

La modelización a la que se arribó tiene la virtud de incorporar explícitamente al sesgo de selección, el que probó ser altamente significativo en el caso de las mujeres de Montevideo y no tanto en el caso de los jóvenes. Para realizar esta prueba se ensayan dos metodologías que ponen de manifiesto elementos que, de no haberlas seguido, hubieran invalidado muchas conclusiones. Así, para el caso de las mujeres, las ecuaciones salariales deberán recoger explícitamente la autoselección si se aspira a contar con estimaciones confiables.

La actividad en la fuerza de trabajo secundaria de Montevideo aparece motivada principalmente por una decisión de optimización de la utilidad familiar y no tanto de utilidades individuales. Para ello se ve influida lógicamente por la conformación demográfica del hogar, en forma agudizada en el caso de las mujeres, de factores asociados al ciclo vital, especialmente en los jóvenes y por factores económicos en ambos grupos. Los elementos individuales, en menor medida, también aportan un grado de explicación a la actividad global.

Es de señalar que el artículo no explica la trayectoria temporal en las tasas de actividad sino que aspira a brindar una explicación de las mismas, una vez aceptada su consolidación. Dicha aspiración brinda elementos para el estudio futuro de la evolución de los determinantes que condujeron a la alta participación de estos grupos en el mercado laboral uruguayo.

Amemiya, T (1981) "Qualitative response models: a survey". Journal of Economic Literature(19):pag.1483-1536.

Diez de Medina, R.(1992) "La estructura ocupacional y los jóvenes en Uruguay". CEPAL Oficina de Montevideo.

Engle, R (1984) "Wald, Likelihood ratio and Lagrange Multiplier tests in Econometrics" en Handbook of Econometrics, Vol II, capítulo 13. North Holland, Amsterdam.

Greene, W (1984) "Estimation of the correlation coefficient in a bivariate probit model using the method of moments", Economics Letters(16):pag.285-291.

_____ (1990) "Econometric Analysis" Macmillan.

_____ (1990) "LIMDEP Manual". Econometric Software, Inc, New York.

Gourieroux, C (1989) "Econometrie des Variables Qualitatives" Economica, Paris.

Heckman, J(1974) "Shadow prices, market wages and labor supply". Econometrica(42): pag.679-694.

_____ (1979) "Sample Selection Bias as a Specification Error" Econometrica(47):pag.153-161.

_____ (1987) "Selection Bias and self-selection" en "Econometrics", The New Palgrave. Ed. Eatwell, J, Milgate, M y Newman, P. Norton & Co. New York.

Kiefer, N (1982) "Testing for independence in Multivariate probit models". Biometrika(69):pag. 161-166.

Killingsworth, M (1983) "Labor supply", Cambridge University Press.

Maddala, G.S (1983) "Limited dependent and Qualitative Variables in Econometrics" ESM, Cambridge University Press.

Mincer, J (1974) "Schooling, Experience and Earnings". NBEA, New York.

Rossi, M. Bucheli, M (1987) "Discriminación laboral contra la mujer" Trabajo presentado a las II Jornadas de Economía del Banco Central del Uruguay.

Rossi, M y Diez de Medina, R (1989) "La mujer en el mercado laboral uruguayo: participación, dedicación, segregación y discriminación". CEPAL Oficina de Montevideo.

_____ (1991) "La actividad femenina en el mercado laboral de Montevideo" Documento 10/91. Departamento de Economía. Facultad de

Ciencias Sociales, Montevideo.

Roy, A.(1950) "The distribution of earnings and of individual output". The Economic Journal (60):pag.489-505, Mayo.

Wynand,P, Bernard,W (1981) "The demand for deductibles in Private health insurance" Journal of Econometrics (17):pag.229-252.