

CENTRO DE INVESTIGACIÓN Y DOCENCIA ECONÓMICAS, A. C.



**EDUCACIÓN E INGRESOS EN MÉXICO: RESULTADOS PARA LA
ENIGH 2002**

T E S I S

**QUE PARA OBTENER EL TÍTULO DE:
LICENCIADO EN ECONOMÍA
PRESENTA:
VICENTE LÓPEZ DÍAZ**

DIRECTOR DE LA TESIS: DR. RODOLFO CERMEÑO BAZAN

MÉXICO, D. F.

Junio 2007

A mi familia

Agradecimientos

Este trabajo no hubiera sido posible sin la guía del Dr. Rodolfo Cermeño. También fueron de mucha ayuda los comentarios de la Dra. Susan Parker, el Dr. John Scott y el Dr. Ángel Salinas. Hubo muchas otras personas que me ayudaron y motivaron a lo largo de este proyecto. A todos ellos muchas gracias.

TABLA DE CONTENIDOS

1. Introducción.....	1
2. Marco Teórico.....	2
3. Estudios empíricos previos	5
4. Datos	6
5. Modelo Econométrico.....	12
6. Estimación Econométrica	14
7. Conclusiones.....	22
BIBLIOGRAFÍA	23

1. Introducción

El objetivo de este estudio es determinar empíricamente la relación entre educación e ingresos. En particular, se considera la necesidad de corregir las estimaciones de retornos a la educación en México por el sesgo de participar o no en la fuerza laboral. La decisión de participar puede arrojar estimadores sesgados, lo que se ha llamado sesgo por selección. Esto es, las personas que perciben algún ingreso pueden tener características muy diferentes a las personas que no perciben ingresos. Los estimadores de los factores que determinan el ingreso pueden ser sobre-estimados si la especificación para estimarlos no toma en cuenta esta posibilidad. El problema de selección puede resultar aún más relevante en caso de existir diferencias de género.

Para evaluar lo anterior se utilizará el procedimiento de selección de Heckman. Por razones prácticas se considera una muestra de la Encuesta Nacional de Ingresos y Gastos de los Hogares de México (ENIGH) del año 2002. La población a analizar son los adultos, mujeres y hombres, con edades entre 25 y 65 años.* El marco teórico utilizado es el modelo de capital humano. Se utiliza también el estimador de mínimos cuadrados ordinarios (MCO) como punto de comparación, pues es un estimador que no considera el problema de selección.

El capítulo 2 presenta el marco teórico. El capítulo 3 muestra una breve revisión de los resultados de estudios empíricos previos. El capítulo 4 contiene una descripción de los datos. El capítulo 5 expone el procedimiento econométrico de Heckman. El capítulo 6 muestra los resultados obtenidos bajo los métodos de estimación MCO y Heckit. Finalmente se presentan las conclusiones.

* No se consideran menores de 25 años por la definición adoptada de experiencia potencial. Ni mayores de 65 por no ser población económicamente activa.

2. Marco Teórico

El presente estudio utiliza el modelo de Mincer (1974), ampliamente utilizado en ecuaciones de ingreso. Lo que sigue es una síntesis del modelo teórico desarrollado por Mincer. La idea central del modelo de capital humano es que el valor presente de los ingresos laborales futuros es igual sin importar los años que se dedicarán a aumentar el nivel de escolaridad. Si el valor presente (VP) de los ingresos futuros por estudiar es mayor al VP obtenido por no estudiar, el individuo, enfrentado a la decisión de estudiar o no, decidirá estudiar. Esto será así hasta que la oferta de personas con mayor educación sea tal que no convenga más seguir estudiando. Para desarrollar esta idea Mincer representó el valor presente del flujo de ingresos de la vida laboral del individuo con escolaridad s (V_s) de la siguiente manera:

$$V_s = Y_s \sum_{t=s+1}^n \left(\frac{1}{1+r} \right)^t \quad (1)$$

Donde n es el tamaño de la vida laboral más *la escolaridad* en años (s), es decir, el tamaño de la vida laboral para las personas sin escolaridad; Y_s es *el ingreso anual* de un individuo con s años de escolaridad; r es *la tasa de descuento* (se supone fija). El tiempo (t), va de 0 hasta n , en años.

Cuando el proceso de descuento es continuo la ecuación equivalente a (1) sería:

$$V_s = Y_s \int_s^n e^{-rt} dt = \frac{Y_s (e^{-rs} - e^{-rn})}{r} \quad (2)$$

Utilizando el resultado anterior y definiendo d como *el diferencial de escolaridad* (entre dos individuos) en años, el valor presente de los ingresos de un individuo por dedicar $s - d$ años de escolaridad es:

$$V_{s-d} = \frac{Y_{s-d}}{r} (e^{-r(s-d)} - e^{-rn}) \quad (3)$$

Con las ecuaciones (2) y (3) se puede construir la razón de ingresos por s años de escolaridad contra $s - d$ años de escolaridad. Esta razón es el valor presente de los ingresos obtenidos por estudiar s años entre el valor presente de estudiar s años menos la diferencia expresada en d . La razón de ingresos tiene las siguientes características: (1) es mayor que uno, más escolaridad significa mayor ingreso; (2) es una función positiva de r , una tasa de descuento más alta significa mayor diferencia en ingresos por año adicional de escolaridad; (3) una función negativa de n , si el horizonte de vida aumenta la diferencia por un año de escolaridad se reduce; (4) la razón de ingresos es una función positiva de s .

Debe notarse que estos efectos son muy pequeños cuando n es grande. Por lo que la razón de ingresos suele tratarse como una constante k . La razón es constante cuando se tiene horizontes de vida laboral fijos, independientemente de la diferencia en años de escolaridad. Lo anterior lleva a escribir la razón como:

$$k_{s,s-d} = \frac{Y_s}{Y_{s-d}} = \frac{e^{-r(s-d)}}{e^{-rs}} = e^{rd} \quad (4)$$

Es decir, la razón sólo depende de la diferencia en escolaridades d no del valor final de la escolaridad s . De la ecuación (3) obteniendo logaritmos se obtiene la ecuación básica del modelo de Mincer:

$$\ln Y_s = \ln Y_0 + rd \quad (5)$$

La ecuación (4) indica que incrementos porcentuales en los ingresos son (estrictamente) proporcionales a las diferencias absolutas del tiempo invertido en la escuela. Es decir, un individuo con dos años más de escolaridad que otro tendrá un ingreso proporcionalmente ma-

yor (por esos dos años). Esa proporción será la misma independientemente de si el individuo con menor escolaridad estudió 5 años o sólo uno.

Mincer completó la ecuación anterior añadiendo la experiencia laboral al modelo. Ante la falta de datos de experiencia laboral efectiva definió la experiencia laboral potencial (edad menos el número de años en la escuela menos 6^{*}) y redefinió la ecuación a estimar como:

$$w = b_0 + b_1s + b_2x + b_3x^2 \quad (6)$$

donde:

w = logaritmo del ingreso laboral

s = nivel de escolaridad

x = experiencia potencial

x^2 = experiencia potencial al cuadrado

Esta es la ecuación que se estima en el presente trabajo.

* Edad de ingreso a la educación primaria.

3. Estudios empíricos previos

Diversos estudios empíricos han estudiado la ecuación de ingreso laboral desde diferentes perspectivas. Gladys López-Acevedo (2001) estimó una ecuación de ingreso laboral utilizando información de la Encuesta Nacional de Empleo Urbano (ENEU). El método de estimación utilizado es la regresión de cuantiles. Se representa la escolaridad en cinco categorías y el valor marginal de cada nivel educativo respecto al previo. Para 1997 el valor marginal de la universidad completa respecto del nivel de preparatoria completa va de 1.63 en el cuantil 10 hasta 1.80 en el cuantil 50.

Otros documentos consideran varios tipos de diferenciales. Mayer y Cordourier (2001) consideran diferenciales por sexo y González (2004) considera diferenciales por status sindical.

Mayer y Cordourier (2001) con información de la ENIGH 1996 estimaron la ecuación de ingreso laboral para dos grupos, hombres y mujeres, utilizaron el método de regresión de cuantiles y obtuvieron resultados para la tasa de retorno de la escolaridad entre 0.078 y 0.137 para hombres y entre 0.095 y 0.106 para mujeres.

González (2004) usa información de la ENIGH 2000 para estimar la ecuación de ingreso laboral considerando diferencias por estatus sindical y sector productivo. Encuentra que la tasa de retorno a la educación es de 0.119 para el sector de servicios y 0.099 para el sector manufacturero entre las personas sindicalizadas.

A diferencia de los estudios mencionados, los cuales también utilizan una ecuación de ingreso laboral similar a la de Mincer (1974), en este trabajo se utiliza el procedimiento de corrección por sesgo de selección de muestra desarrollado por Heckman (1974).

4. Datos

La base de datos que se utiliza en el presente trabajo es la Encuesta Nacional de Ingresos y Gastos de los Hogares de 2002 (ENIGH 2002). Esta encuesta proporciona información sobre características sociodemográficas de la población mexicana, pues es de representatividad nacional, incluyendo zonas rurales y urbanas. Existe información de ingresos a nivel individual, razón que influyó en la decisión de utilizar esta base de datos.

La muestra analizada incluye sólo a los miembros del hogar con edades entre 25 y 65 años de edad, con algún tipo de ingreso laboral (es decir, con remuneraciones por trabajo, según el catálogo de la ENIGH).

La variable de interés es el *ingreso por hora*. Es conveniente utilizar el salario por hora para obtener una medida de ingreso laboral que sea independiente del número de horas dedicadas al trabajo. Si la variable ingreso fuera semanal, mensual o anual esto no ocurriría. Se incluye en la muestra sólo a las personas con uno o dos empleos. En este trabajo se analiza el ingreso laboral en la muestra por sexo, nivel educativo y grupo de edad.

De acuerdo con la ENIGH, la población total en 2002 era de 101,522,414 personas, de las cuales 49,459,117 (48.72 %) eran hombres y el resto 52,063,297 (41.28 %) eran mujeres. En el grupo de edad entre 25 y 65 años había 20,277,685 (46.21 %) hombres y 23,601,556 (53.79 %) mujeres, en total 43,879,241 individuos. Entre los hombres 13,964,908 son jefes de familia (con pareja) con uno o dos empleos y 6, 182, 680 son cónyuges con uno o dos empleos. En el caso en que tanto el jefe de familia como el cónyuge tengan ingreso por remuneraciones al trabajo se considera sólo la información del jefe de familia. Este es el grupo de población de interés. Esta población se subdivide en dos grupos, mostrados en las tablas 4.a y

4.b, los que declararon percibir un ingreso en el mes previo a la entrevista y los que declararon no percibir ingreso alguno.

Tabla 4.a			
Población considerada en las estimaciones			
	hombre	mujer	Total
sin ingresos	7,329,009	16,758,002	24,087,011
con ingresos	12,948,676	6,843,554	19,792,230
Total	20,277,685	23,601,556	43,879,241

En la tabla 4.a se muestran los estimados poblacionales y en la tabla 4.b los observados en la muestra. Se presenta la distribución de las variables de interés, la escolaridad y la edad, para cada uno de estos grupos. Además se estudia la distribución de los ingresos laborales donde es posible.

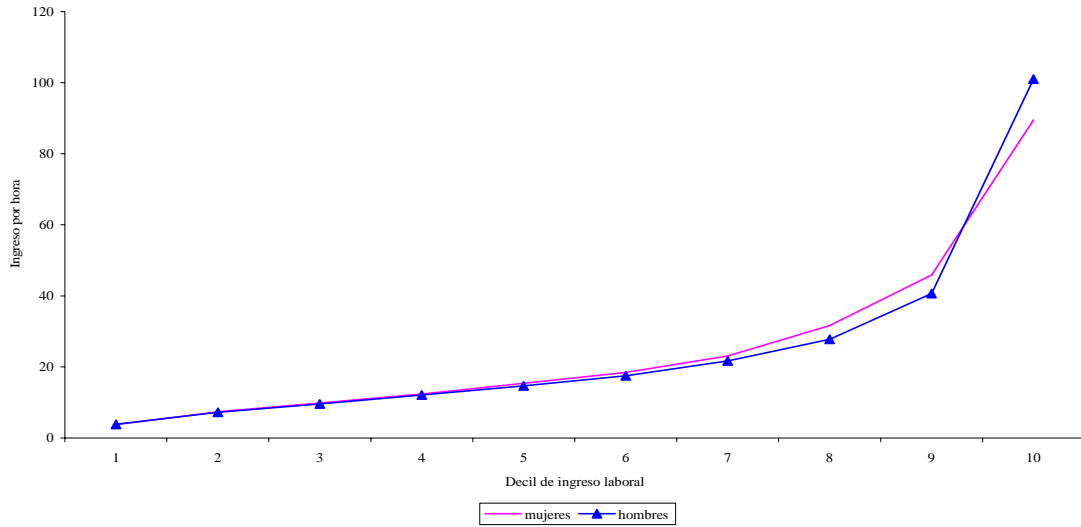
Tabla 4.b			
Población considerada en las estimaciones (muestra)			
	hombre	mujer	Total
sin ingresos	5,110	11,722	16,832
con ingresos	8,966	4,463	13,429
Total	14,076	16,185	30,261

Lo anterior se hace con la intención de orientar la búsqueda de un modelo de ingresos que mejor se aplique a los datos observados. Se busca identificar diferencias en la distribución de las variables explicativas en dos niveles: por género y por percepción de ingresos.

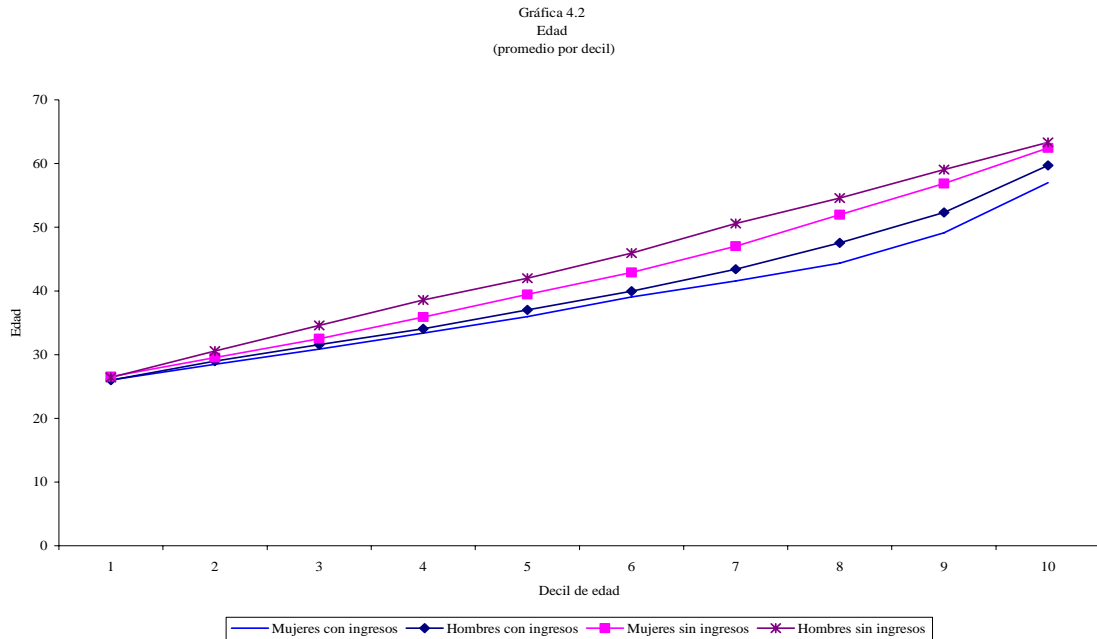
En la Gráfica 4.1 se muestra el promedio del ingreso (por hora) por decil de ingreso para hombres y mujeres. Ahí se observa que si se ordena por ingreso a las mujeres y a los hombres y se les compara uno a uno, el promedio del ingreso de las mujeres que están entre el

70 y el 90 por ciento con más ingresos reciben mayores percepciones, en promedio, que los hombres en el mismo decil. En el último decil la situación se revierte, son los hombres los que ganan más en promedio que las mujeres. Del decil uno al seis no hay un orden evidente.

Gráfica 4.1
Ingreso por hora
(promedio por decil)



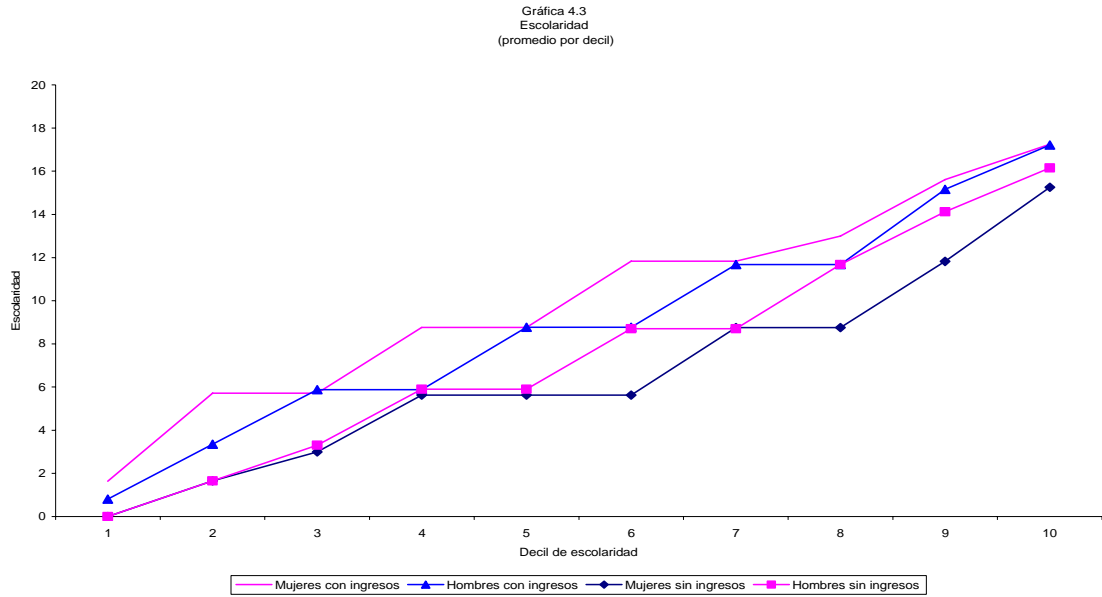
Si se considera la distribución de las edades también se observan diferencias entre hombres y mujeres que pueden explicar posibles diferencias en la percepción de ingresos. En la gráfica 4.2 se observa que la edad es en promedio mayor entre los hombres que entre las mujeres. Esto es, entre los hombres que perciben ingresos se observa una distribución de edades más alta a la distribución de edades de las mujeres que perciben ingresos.



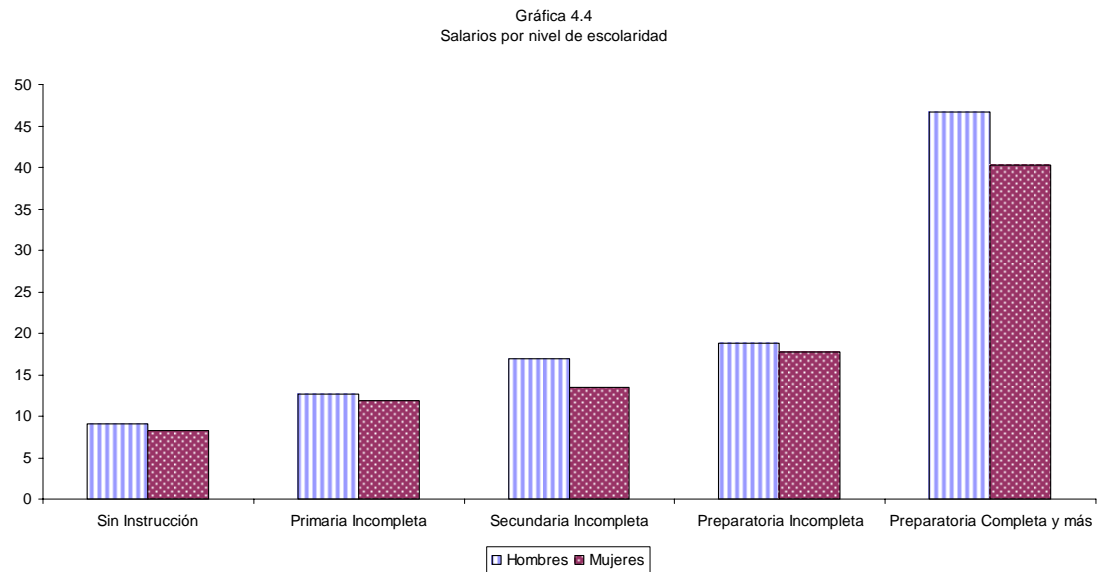
En la gráfica 4.2 también se muestra la edad promedio entre quienes no reportaron ningún ingreso. En promedio se trata de personas mayores. Las mujeres que no perciben ingresos son mayores a los hombres que si los perciben y los hombres que no perciben ingresos son en promedio mayores a las mujeres en la misma situación.

La situación inversa se observa en la distribución de los años de escolaridad. Esto se muestra en la gráfica 4.3. Las mujeres que perciben ingresos tienen promedios de escolaridad mayores a los equivalentes en los hombres en la misma condición de percepción.

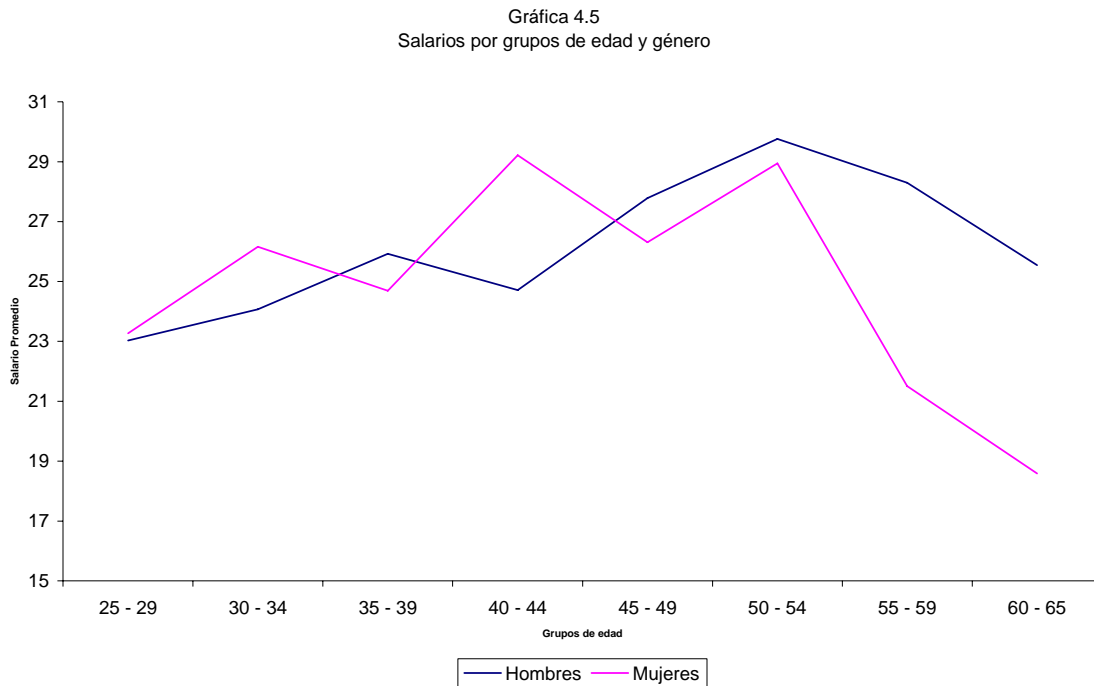
En la gráfica 4.3 se presenta el número promedio de años dedicados a la escuela para los que no perciben ningún ingreso. Se observa el patrón inverso al que se observa en la edad. Esto puede indicar que las personas con mayor edad tienen menor escolaridad. Sin embargo, para la estimación del modelo de ingresos del siguiente capítulo es importante tener estas diferencias en cuenta, pueden explicar los resultados del modelo.



En la gráfica 4.4 se muestra el salario por hora promedio de hombres y mujeres por nivel de escolaridad. En todos los casos se observa un salario menor para las mujeres. En esta gráfica se puede observar, además, que la diferencia en el ingreso promedio por hora entre hombres y mujeres aumenta considerablemente en el nivel de escolaridad más alto.



En la gráfica 4.5 se muestra igualmente el promedio del ingreso laboral dividiendo la población por grupos de edad (cada grupo incluye cinco años). Se observa que para algunos grupos de edad el ingreso laboral de las mujeres es mayor al de los hombres Pero en la mayoría de los casos, cinco de ocho, se observa que el ingreso laboral promedio de los hombres es mayor al de las mujeres. Se advierte que a partir de 45 años los individuos de sexo masculino tienen siempre empleos mejor remunerados que las mujeres Ambos grupos sufren una caída después de los 54 años, misma que es más dramática entre las mujeres que entre los hombres. El ingreso laboral promedio muestra mayores variaciones entre las mujeres que entre los hombres.



5. Modelo Econométrico

La estimación general de un modelo como el presentado anteriormente podría ajustarse a los supuestos del modelo clásico de regresión lineal. Sin embargo, la posibilidad de auto-selección en la muestra elegida obliga a que se consideren otros métodos de estimación. Entre los modelos que han sido desarrollados para este tipo de problemas se encuentra el modelo de Heckman (1979). Como se expone en Johnston & DiNardo (1997) el modelo a estimar sería el siguiente:

$$w_i = X_i\beta + \varepsilon_{1i} \quad (3.1)$$

donde w es el logaritmo del salario, X es un vector de características que influyen en la determinación del salario, en este caso (*edad, edad al cuadrado y escolaridad*). Como el resultado de la ecuación anterior puede ser afectado por selección de muestra es necesario considerar esta situación para obtener estimadores insesgados. Para ello se puede considerar la ecuación de participación:

$$T_i = 1(Z_i\gamma + \varepsilon_{0i} > 0) \quad (3.2)$$

donde Z es un vector con las variables que explican la participación laboral. En este caso además de las variables que explican el nivel del salario se consideran *el número de niños pequeños en el hogar*, entre otras características del individuo y del hogar (o de la vivienda). Si los términos de error en ambas ecuaciones están correlacionados existe un sesgo por selección de muestra. Este sesgo puede ser considerado como si se omitiera una variable.

Para corregir el sesgo, si la distribución de los errores es normal conjunta, se puede utilizar el siguiente procedimiento³:

1. Estimar un modelo del tratamiento (3.2) mediante un probit.
2. Usar los estimados para construir la razón inversa de Mills.
3. Estimar el modelo de ingreso en las variables de X añadiendo la razón inversa de Mills como regresor.

³ La regresión a estimar en el paso (3) se especifica como

$$E[y_i | X_i, T_i = 1] = X_i\beta + \frac{\phi(Z_i\gamma / \sigma_0)}{\Phi(Z_i\gamma / \sigma_0)} \tilde{\sigma}$$

En el segundo término del lado derecho aparece la razón inversa de Mills. Para implementar la regresión anterior se utiliza un estimado obtenido de un modelo probit tal como se indica en los pasos (1) y (2).

6. Estimación Econométrica

En este capítulo se presenta la estimación econométrica del modelo de capital humano que se desarrolló en los capítulos anteriores. En la tabla 6.1 se incluyen las estadísticas descriptivas de las variables utilizadas en el modelo. En las siguientes tablas, 6.2 a la 6.4, se muestran los resultados de la estimación con diferentes especificaciones.

En la tabla 6.1 se observan algunos de los detalles ya mencionados en el capítulo 4. Las personas con mayor *escolaridad* y menor *edad* son las que reportan tener ingresos. Esto sugiere la necesidad de estimar el modelo con corrección por el sesgo de selección. En la tabla 6.2 se muestran los resultados de la estimación por mínimos cuadrados ordinarios. La desviación estándar de los estimadores de los coeficientes ha sido corregida por la posible presencia de heteroscedasticidad.

Tabla 6.1							
Promedios de variables en el modelo							
Variable	Nacional	Con ingresos			Sin ingresos		
		Total	Mujeres	Hombres	Total	Mujeres	Hombres
Logaritmo del ingreso laboral por hora	-	2.82	2.84	2.80	-	-	-
Escolaridad	7.55	8.78	9.44	8.44	6.53	6.28	7.10
Edad	41.01	38.94	37.98	39.45	42.71	42.06	44.19
Nivel de instrucción							
Sin Instrucción	9.86	5.68	4.14	6.5	13.3	13.91	11.89
Primaria Incompleta	20.15	15.61	13.33	16.82	23.88	24.29	22.96
Secundaria Incompleta	24.5	21.87	20.16	22.78	26.66	27.39	24.99
Preparatoria Incompleta	18.57	20.88	17.95	22.43	16.66	16.72	16.54
Preparatoria Completa y más	26.92	35.96	44.43	31.48	19.5	17.69	23.63
Grupo de edad							
25 - 29	17.48	20.06	21.19	19.47	15.36	16.26	13.3
30 - 34	16.53	18.8	19.86	18.24	14.67	15.62	12.48
35 - 39	15.86	18.02	18.74	17.64	14.09	14.21	13.82
40 - 44	14.58	15.45	17.15	14.54	13.87	14.33	12.84
45 - 49	10.95	10.89	9.44	11.65	10.99	10.89	11.23
50 - 54	9.54	8.02	7.68	8.2	10.79	10.45	11.57
55 - 59	7.52	5.03	3.4	5.89	9.57	9.1	10.67
60 - 65	7.53	3.73	2.54	4.36	10.65	9.15	14.08

Tabla 6.2						
Variable dependiente: Logaritmo del ingreso laboral por hora						
Variable explicativa	(1)	(2)	(3)	(4)	(5)	(6)*
Escolaridad	0.139 (44.16)***	0.14 (45.09)***	0.174 (20.48)***	0.167 (18.52)***	0.163 (17.86)***	0.166 (19.83)***
Experiencia potencial	0.013 (11.63)***	0.033 (8.55)***	0.063 (7.94)***	0.064 (7.50)***	0.063 (7.06)***	0.066 (8.18)***
Experiencia potencial al cuadrado		-0.0003 (5.42)***	-0.001 (6.57)***	-0.001 (6.31)***	-0.001 (5.85)***	-0.001 (6.78)***
Experiencia potencial por escolaridad			-0.001 (3.91)***	-0.001 (3.67)***	-0.001 (3.51)***	-0.002 (4.45)***
Comunidad rural con menos de 2500 habitantes				-0.387 (10.12)***	-0.429 (10.95)***	-0.322 (8.55)***
Comunidad con entre 2500 y 15 mil habitantes					-0.243 (6.75)***	-0.145 (4.51)***
Constante	1.276 (25.35)***	1.05 (16.17)***	0.518 (3.98)***	0.646 (4.78)***	0.723 (5.22)***	0.695 (5.39)***
Observaciones	13429	13429	13429	13429	13429	13429
Coefficiente de determinación	0.4	0.4	0.4	0.43	0.43	0.47

Valor absoluto del estadístico t en paréntesis

* significativo al 10%; ** significativo al 5%; *** significativo al 1%

La tabla 6.2 muestra los resultados de la estimación de la ecuación de ingresos incluyendo en la muestra tanto a hombres como a mujeres. Se observa cómo la inclusión de diferentes regresores afecta la estimación de la tasa de rendimientos a la educación.

Como se esperaba, el ingreso laboral muestra rendimientos decrecientes con respecto a la experiencia potencial. Lo cual significa que el aumento en la experiencia potencial aumenta el nivel esperado de ingreso laboral pero a tasas decrecientes. Un factor que disminuye el ingreso laboral esperado es el vivir en una comunidad rural. En la especificación 6 se incluyeron efectos fijos a nivel de entidad federativa.

En las tablas 6.3 y 6.4 se plantean las mismas especificaciones que en la tabla 6.2. En la tabla 6.3 sólo se considera la muestra de mujeres y en la tabla 6.4 sólo se considera la muestra de hombres. En estas dos especificaciones se observa entre otras cosas que los rendimientos a la escolaridad son mayores entre las mujeres que entre los hombres. Además entre las mujeres, la *experiencia potencial* muestran mayor influencia en el ingreso laboral. Esto significa que al contratar un hombre la experiencia no es tan valorada como cuando se contrata una mujer.

Tabla 6.3						
Variable dependiente: Logaritmo del ingreso laboral por hora (mujeres)						
Variable explicativa	(1)	(2)	(3)	(4)	(5)	(6)
Escolaridad	0.147 (24.82) ^{***}	0.148 (50.23) ^{***}	0.191 (13.67) ^{***}	0.191 (13.80) ^{***}	0.187 (12.97) ^{***}	0.192 (15.30) ^{***}
Experiencia potencial	0.015 (7.86) ^{***}	0.031 (8.58) ^{***}	0.069 (4.58) ^{***}	0.073 (4.96) ^{***}	0.071 (4.35) ^{***}	0.077 (5.86) ^{***}
Experiencia potencial al cuadrado		-0.0002 (4.57) ^{***}	-0.001 (3.43) ^{***}	-0.001 (3.78) ^{***}	-0.001 (3.22) ^{***}	-0.001 (4.62) ^{***}
Experiencia potencial por escolaridad			-0.002 (2.97) ^{***}	-0.002 (3.32) ^{***}	-0.002 (2.96) ^{***}	-0.002 (4.06) ^{***}
Comunidad rural con menos de 2500 habitantes				-0.353 (4.21) ^{***}	-0.376 (4.47) ^{***}	-0.3 (3.80) ^{***}
Comunidad con entre 2500 y 15 mil habitantes					-0.151 (1.89) [*]	-0.102 (1.72) [*]
Constante	1.11 (11.74) ^{***}	0.939 (15.03) ^{***}	0.269 -1.16	0.285 -1.25	0.355 -1.46	0.347 -1.63
Observaciones	4463	4463	4463	4463	4463	4463
Coefficiente de determinación	0.41	0.41	0.41	0.43	0.43	0.47

Valor absoluto del estadístico t en paréntesis

* significativo al 10%; ** significativo al 5%; *** significativo al 1%

Tabla 6.4						
Variable dependiente: Logaritmo del ingreso laboral por hora (hombres)						
Variable explicativa	(1)	(2)	(3)	(4)	(5)	(6)*
Escolaridad	0.136 (35.78)***	0.137 (36.28)***	0.168 (15.59)***	0.155 (13.52)***	0.152 (13.26)***	0.155 (14.63)***
Experiencia potencial	0.012 (8.71)***	0.034 (6.95)***	0.061 (6.55)***	0.061 (5.78)***	0.061 (5.72)***	0.063 (6.43)***
Experiencia potencial al cuadrado		-0.0003 (4.84)***	-0.001 (5.67)***	-0.001 (5.08)***	-0.001 (5.04)***	-0.001 (5.52)***
Experiencia potencial por escolaridad			-0.001 (2.82)***	-0.001 (2.43)**	-0.001 (2.44)**	-0.001 (3.17)***
Comunidad rural con menos de 2500 habitantes				-0.43 (10.01)***	-0.484 (10.78)***	-0.363 (8.26)***
Comunidad con entre 2500 y 15 mil habitantes					-0.296 (8.21)***	-0.177 (5.05)***
Constante	1.351 (22.53)***	1.093 (13.17)***	0.613 (3.88)***	0.813 (4.87)***	0.889 (5.36)***	0.829 (5.28)***
Observaciones	8966	8966	8966	8966	8966	8966
Coefficiente de determinación	0.39	0.4	0.4	0.43	0.44	0.49

Valor absoluto del estadístico t en paréntesis

* significativo al 10%; ** significativo al 5%; *** significativo al 1%

Para considerar el efecto de auto-selección se propone como variable instrumental el número de niños en el hogar.

En la tabla 6.5 se incluye la estimación del mismo modelo corrigiendo por el sesgo de selección de la muestra. Se muestran los resultados de la ecuación de participación laboral (selección) así como los resultados de la ecuación de ingresos. Las variables que sólo se consideraron en la ecuación de participación laboral son del hogar (presencia y número de menores). Si la estimación por este procedimiento es correcta, se observa que el retorno a la educación está sub-estimado en el modelo de Mínimos Cuadrados. Sin embargo en cada caso la razón de Mills indica que no existe sesgo por selección de muestra. Estos resultados muestran que los indicios de sesgo de auto-selección, señalados en capítulo 4, no son suficientes para justificar una corrección en los estimadores. Tal vez los instrumentos elegidos no sean los más adecuados para este tipo de estimación.

A diferencia de los estudios para el caso mexicano, que se revisaron en el capítulo 3, los retornos a la educación estimados en este trabajo son mayores. La determinación de las causas específicas puede ser motivo de otra investigación. En especial pueden explorarse los cambios en restricciones de crédito a la educación que pueden afectar la decisión de continuar estudiando. Además pueden explorarse otros instrumentos en la ecuación de selección.

Tabla 6.5						
Variable dependiente: Logaritmo del ingreso laboral por hora						
Método de estimación: Regresión con el procedimiento de Heckman						
Variable explicativa	Población		Mujeres		Hombres	
	Salario	Selección	Salario	Selección	Salario	Selección
Escolaridad	0.145 (13.72) ^{***}	0.08 (7.19) ^{***}	0.185 (12.98) ^{***}	0.133 (6.90) ^{***}	0.155 (14.57) ^{***}	0.012 -0.85
Experiencia potencial	0.053 (6.01) ^{***}	0.048 (4.66) ^{***}	0.074 (5.50) ^{***}	0.063 (3.08) ^{***}	0.064 (6.41) ^{***}	0.028 (2.26) ^{**}
Experiencia potencial al cuadrado	-0.001 (3.88) ^{***}	-0.001 (6.03) ^{***}	-0.001 (4.25) ^{***}	-0.001 (3.50) ^{***}	-0.001 (5.53) ^{***}	-0.001 (4.46) ^{***}
Experiencia potencial por escolaridad	-0.001 (3.12) ^{***}	-0.002 (4.32) ^{***}	-0.002 (3.68) ^{***}	-0.003 (4.33) ^{***}	-0.001 (3.17) ^{***}	-0.001 -1.45
Comunidad rural con menos de 2500 habitantes	-0.265 (6.40) ^{***}	-0.164 (4.61) ^{***}	-0.277 (3.41) ^{***}	-0.427 (7.95) ^{***}	-0.366 (8.23) ^{***}	-0.155 (2.97) ^{***}
Comunidad con entre 2500 y 15 mil habitantes	-0.116 (3.29) ^{***}	-0.091 (1.82) [*]	-0.099 (1.69) [*]	-0.047 -0.59	-0.18 (5.04) ^{***}	-0.185 (2.75) ^{***}
Número de niños entre 0 y 5 años		-0.016 -1.05		-0.169 (6.36) ^{***}		0.069 (2.85) ^{***}
Número de niños entre 6 y 12 años		-0.056 (4.57) ^{***}		-0.092 (4.54) ^{***}		-0.009 -0.44
Número de niños entre 13 y 18 años		0.0002 -0.02		0.023 -0.96		-0.004 -0.17
Constante	1.252 (5.69) ^{***}	-0.81 (4.65) ^{***}	0.494 (1.90) [*]	-1.508 (4.61) ^{***}	0.813 (5.03) ^{***}	0.386 (1.69) [*]
Observaciones	30261	30261	16185	16185	14076	14076

Valor absoluto del estadístico z robusto en paréntesis

* significativo al 10%; ** significativo al 5%; *** significativo al 1%

7. Conclusiones

Las diferentes especificaciones de la ecuación de salario estimadas en este trabajo permiten inferir que el mercado de trabajo no paga de la misma forma el nivel de escolaridad entre hombres y mujeres. Al estimar la ecuación para una muestra restringida a las personas con ingresos la tasa de retornos a la educación es mayor para mujeres que para hombres. Sin embargo, si se considera el procedimiento de corrección por sesgo de selección de Heckman el orden cambia. De lo anterior se infiere que entre las mujeres que no perciben ingresos el retorno por cada año de escolaridad debe ser mayor para que compense las diferentes actividades que dejan de hacer por trabajar por un ingreso. Esto es, existen, por lo menos en el mercado laboral de mujeres, restricciones que afectan la decisión de trabajar. Esto se observa en las estimaciones del rendimiento de la escolaridad cuando se contrasta la estimación sin considerar la decisión de participación en el mercado laboral contra la estimación considerando la misma decisión.

El presente trabajo es concluyente respecto a la necesidad de considerar la selección en el mercado laboral de las mujeres para México. En estudios posteriores se podría explorar si esto es así recientemente, si la diferencia ha disminuido en el transcurso del tiempo y qué factores podría explicar estos cambios.

BIBLIOGRAFÍA

- Ashenfelter, O. and Heckman, James, “The Estimation of Income and Substitution Effects in a Model of Family Labor Supply” en *Econometrica*, Enero 1974, 42(1), pp.73-86.
- Berndt, Ernest. *The practice of Econometrics*. Addison-Wesley, 1991.
- Greene, William H., *Econometric Analysis*, 5ed, Prentice Hall, 2003.
- González Ibargüen, Juan Carlos, *Análisis de diferenciales de salario por status sindical*, Tesis, Centro de Investigación y Docencia Económicas, 2004.
- Heckman, James, “Shadow Prices, Market Wages and Labor Supply” en *Econometrica*, Julio 1974, 42(1), pp.679-94.
- Instituto Nacional de Estadística Geografía e Informática, *Encuesta Nacional de Ingresos y Gastos de los Hogares 2002*, México, 2003.
- Jonhston, Jack y John DiNardo, *Econometric Methods*, 4ed, McGraw-Hill, 1997.
- López-Acevedo, Gladys, *Evolution of Earnings and Rates of Returns to Education in Mexico*, World Bank, Policy Research Working Paper 2691, 2001.
- Mayer Foulkes, David y Gabriela Cordourier Real, “La brecha salarial y la teoría de igualdad de oportunidades” en *El Trimestre Económico*, vol. LXVIII(1), num. 269, pp.71-107, 2001.
- Mincer, Jacob. *Schooling, Experience and Earnings*, National Bureau of Economic Research, Nueva York, 1974.