División de Economía

Número 44

Andrés Zamudio Carrillo

Rendimientos a la educación superior en México: ajuste por sesgo utilizando máxima verosimilitud

Las colecciones de Documentos de Trabajo del CIDE representan un medio para difundir los avances de la labor de investigación, y para permitir que los autores reciban comentarios antes de su publicación definitiva. Se agradecerá que los comentarios se hagan llegar directamente al(los) autor(es). ❖ Los editores han mantenido fielmente el texto original del presente documento de trabajo, por lo que tanto el contenido como el estilo y la redacción son responsabilidad exclusiva del(de los) autor(es). Coordinadora de Publicaciones: María Ofelia Arruti H. ❖ Producción: Solar, Servicios Editoriales, S.A. de C.V., Calle 2, núm. 21, San Pedro de los Pinos, 03820 México, D.F., tels. 515-1657 y 271-9027

Introducción

Los rendimientos a la educación, y en particular a la educación superior, ha sido un tema poco estudiado en México. El retorno o rendimiento es uno de los elementos importantes en la determinación de la demanda de servicios educativos por parte de los individuos. Asimismo, los retornos son importantes para la política económica. Retornos relativamente altos o bajos en algún sector o ciclo escolar pueden ayudar a identificar posibles situaciones de escasez relativa de personal calificado.

Existen pocas estimaciones sobre los rendimientos a la educación en México. Como ejemplos de estos trabajos se encuentran M. Carnoy (1967), G. Psacharopoulos y Y. Chu Ng (1992); T. Bracho, y A. Zamudio (1994), y A. Zamudio y T. Bracho (1994). En todos ellos, a diferencia del último estudio, se lleva a cabo una estimación de los rendimientos o retornos sin tomar en cuenta el posible sesgo que resulta de no considerar a la escolaridad como una variable de elección del individuo.[1]

En A. Zamudio y T. Bracho (1994) se estiman retornos a la educación libres del problema de sesgo por elección, pero estos retornos se calculan para la educación en general, sin considerar las particularidades que pueda tener cada uno de los ciclos educativos.

En el presente documento se trabaja con un segmento del ciclo escolar: la educación superior. Para este sector reviste especial importancia el problema de la endogeneidad de la escolaridad, ya que en el acceso a la educación superior en México influyen tanto las características de los individuos como los factores socioeconómicos. Muchos de estos factores forman parte, al mismo tiempo, de la decisión de continuar con estudios superiores y del desempeño laboral del individuo. A causa de que muchos de estos factores no son observables, o son de difícil medición, existe el riesgo de obtener estimadores sesgados de los retornos.

Para llevar a cabo el cálculo de los retornos se estiman ecuaciones de ingreso y de elección del individuo. Los parámetros de las ecuaciones de ingreso, una vez libres del problema de sesgo, son utilizados para calcular los retornos o rendimientos.

El método utilizado en la estimación es el de máxima verosimilitud, aunque también se incluyen, con el propósito de hacer comparaciones, estimaciones basadas en mínimos cuadrados ordinarios y el método bietápico de Heckman o “Heckit”.[2]

Algunos estudios previos sobre la estimación de los retornos a la educación superior con ajuste por sesgo se encuentran en L. Kenny, L. Lee, G. Maddala y R. Trost (1979), quienes utilizan máxima verosimilitud; E. Cohn y W. Hughes (1994) o R. Willis y S. Rosen (1979), quienes utilizan métodos bietápicos.

[1] Una discusión sobre el problema de sesgo por elección se encuentra en A. Zamudio y T. Bracho (1994).

[2] Véase J. Heckman (1979).

Método de estimación[3]

Para llevar a cabo el cálculo de los rendimientos se estiman, primero, ecuaciones de ingreso para individuos con educación media (preparatoria) y para los que tienen estudios superiores. Una vez obtenidos los parámetros de las dos ecuaciones de ingreso, los cuales deben estar libres del problema de sesgo por elección, los rendimientos se calculan al obtener la tasa interna de retorno, que iguale el valor presente de los ingresos futuros para ambos proyectos educativos.

Denotemos con el subíndice "0" a la primera situación, es decir, educación media, y con "1" a la segunda: continuar con estudios superiores.

La forma funcional para las ecuaciones de ingreso será del tipo semilogarítmica, esto es, el ingreso en forma logarítmica y las variables explicativas en forma lineal.[4] El ingreso del individuo depende del nivel de escolaridad, de la experiencia laboral y de ciertas características personales o familiares observables del individuo (denotadas por el vector X_i), y de ciertas características no observables (u_{Si}). La función de ingresos será denotada por $Y(S, X_i)$, donde S asume solamente dos valores: 0 y 1. Si las variables explicativas afectan de diferente manera la formación de ingresos para cada situación o régimen, entonces podemos escribir una ecuación de ingresos para cada régimen. De este modo las expresiones quedarían como:

$$\ln[Y(S = 0, X_i)] = y_{0i} = X_i' \beta_0 + u_{0i} \tag{1}$$

$$\ln[Y(S = 1, X_i)] = y_{1i} = X_i' \beta_1 + u_{1i}, \tag{2}$$

donde β_0 y β_1 corresponden a los vectores de parámetros de interés.

Las expresiones (1) y (2) forman un sistema de ecuaciones de dos regímenes (*switching model*), donde ambas ecuaciones no son observables simultáneamente para un cierto individuo. Si el individuo decide terminar sus estudios en el nivel medio, entonces se observa la ecuación (1); de otro modo, la ecuación (2).

La estimación simple por mínimos cuadrados ordinarios de (1) y (2) puede dar por resultado estimadores sesgados de los parámetros. Este posible sesgo es producto de no considerar a la decisión de educación (superior en este caso) como una variable endógena.[5]

Si la educación es una variable endógena, entonces (1) o (2) será observado, dependiendo de la elección del individuo.

La educación puede ser vista como una inversión o como un consumo. De esta manera, variables como la educación de los padres o la zona de residencia pueden ser

[3] La presente discusión sobre el método de estimación es conocida en la literatura; véase, por ejemplo, L. Lee (1979) o G. Maddala (1983).

[4] Esta forma funcional para la ecuación de ingresos es común en este tipo de estudios y es similar a la llamada ecuación Minceriana.

[5] Véase J. Heckman (1979) o G. Maddala (1983).

tan importantes como el costo de la educación o el valor presente de los ingresos esperados en el futuro. Supondremos que existe un conjunto de factores exógenos que explica la decisión de educación.[6]

Estos factores consisten tanto en variables observables como no observables o medibles. Denotemos estos observables por medio del vector W_i y los no observables mediante la variable aleatoria e_i, y definamos, así, la variable latente I_i^* de la siguiente manera:

$$I_i^* = W_i'\alpha - e_i. \tag{3}$$

La variable I_i^* no es observable en forma continua, solamente es observable cuando toma valores positivos o negativos. Supongamos que cuando el individuo elige llevar a cabo estudios superiores, la variable latente I_i^* toma valores positivos, mientras que en el caso contrario, el valor es negativo o cero. Definamos a la variable dicotómica I_i como:

$$\begin{aligned} I_i &= 1 \ \text{ cuando } I_i^* > 0, \\ I_i &= 0 \ \text{ de otro modo.} \end{aligned} \tag{4}$$

Asumiendo normalidad en los residuales e_i se llega a un modelo tipo probit sobre la elección de continuar o no con la educación superior. Como (2) es observado cuando $I_i^* > 0$, el sesgo por elección es producto de la covarianza entre el residual de (2) y el residual de (3), es decir, (2) es observado cuando $W_i'\alpha > e_i$. Del mismo modo, (1) es observado cuando $W_i'\alpha \leq e_i$.

Supongamos que el vector de variables aleatorias (u_{0i}, u_{1i}, e_i) tiene una distribución normal trivariada, con esperanzas incondicionales cero y matriz de varianzas y covarianzas incondicionales,

$$\Sigma = \begin{pmatrix} \sigma_0^2 & \sigma_{01} & \sigma_{0e} \\ \sigma_{01} & \sigma_1^2 & \sigma_{1e} \\ \sigma_{0e} & \sigma_{1e} & \sigma_e^2 \end{pmatrix}. \tag{5}$$

Para esta matriz de covarianzas tenemos las siguientes restricciones. Puesto que se tiene un modelo de dos regímenes, el parámetro σ_{01} no es estimable al no ser observables simultáneamente las primeras dos ecuaciones (ecuaciones de ingreso).

[6] En algunos estudios, como en el de R. Willis y S. Rosen (1979), o en el de L. Lee (1978), se utilizan las ecuaciones de ingreso para obtener una proxy de los retornos esperados. Sin embargo, este procedimiento tiene sus problemas para el presente caso. Al utilizar información de corte transversal se está suponiendo que los ingresos actuales son equivalentes a los ingresos esperados en el momento de tomar la decisión. Obviamente, éste no sería el caso cuando se tiene una situación de cambio. Por esta razón, no se incluyeron las ecuaciones de ingreso en la ecuación de elección (lo que daría lugar a un sistema de ecuaciones simultáneas). La elección se hizo, en cambio, en función de variables exógenas.

Como la tercera ecuación es del tipo probit, la varianza para e_i no es identificable, por lo cual se asume que esta varianza es igual a la unidad.

Las probabilidades de observar cada uno de los dos regímenes se pueden expresar como:

$$P[Y_{1i} \text{ sea observado}]=P[I_i^* > 0] = P[W_i'\alpha > e_i] = \Phi(W_i'\alpha)$$
$$P[Y_{0i} \text{ sea observado}] = P[I_i^* \leq 0] = P[W_i'\alpha \leq e_i] = 1 - \Phi(W_i'\alpha),$$

donde $\Phi(W_i'\alpha)$ representa la función de distribución normal estandarizada evaluada en el punto $W_i'\alpha$. En este caso, la probabilidad de que el régimen "1" sea observado va a estar en función directa al valor del argumento $W_i'\alpha$, mientras que la probabilidad de observar el régimen "0" estará en función inversa al argumento.

Utilizando resultados conocidos para distribuciones normales truncadas[7] la posibilidad de sesgo por elección para el régimen "0" se puede ver en las siguientes expresiones:

$$\begin{aligned} &E[Y_{0i} \mid \text{ régimen "0" es observado}] = E[Y_{0i} \mid W_i'\alpha \leq e_i] = \\ &X_i'\beta_0 + E[u_{0i} \mid W_i'\alpha \leq e_i] = X_i'\beta_0 + \sigma_{0e}\, \phi_i/(1 - \Phi_i). \end{aligned} \tag{6}$$

ϕ_i y Φ_i corresponden a las funciones de densidad y distribución normal estandarizada evaluadas en el punto $W_i'\alpha$. El sesgo en las estimaciones de los parámetros β surge de la segunda expresión, ya que la media de los residuales, dado el proceso de selección, no es cero.

Cuando el valor del parámetro σ_{0e} es cero no se tiene problema de sesgo. Sin embargo, cuando éste no es el caso, se pueden presentar las siguientes situaciones.

Puede demostrarse que la expresión $\phi_i/(1 - \Phi_i)$ es una función creciente del argumento $W_i'\alpha$,[8] la cual a su vez se encuentra en función indirecta respecto de la probabilidad de inclusión en la muestra. De este modo, si para el individuo i la probabilidad de ser observado en el régimen "0" es muy alta, entonces la expresión $\sigma_{0e}\, \phi_i/(1 - \Phi_i)$ tenderá a ser muy pequeña, por lo que el problema de sesgo por elección sería despreciable. En el otro caso, cuando la probabilidad de ser observado en "0" es muy pequeña, lo cual indica que la probabilidad de ser observado en "1" es muy alta, el problema de sesgo puede ser muy importante.

En resumen, la posibilidad de sesgo por elección en la estimación de la ecuación de ingresos para un cierto régimen (en este caso el régimen "0"), va a ser mayor mientras más baja sea la probabilidad de encontrarse en él.

Supongamos que para el individuo i el régimen "0" es observado. Cuando el modelo probit "pronostica" mal, esto es, predice que el régimen "1" es observado cuando en realidad no lo es, el valor de e_i resulta mayor a $W_i'\alpha$, y puesto que esta última

[7] Véase, por ejemplo, L. Lee y R. Trost (1978) o G. Maddala (1983).

[8] Véase L. Lee (1979).

expresión tiene que ser positiva, entonces e_i es positivo. Si el parámetro σ_{0e} es positivo, entonces se esperaría que valores positivos de e_i estuvieran asociados con valores positivos de u_{0i}, por lo que el valor esperado de u_{0i} es mayor que cero.[9]

Para la ecuación de ingresos correspondiente a la educación superior tenemos la siguiente expresión:

$$\begin{aligned} E[Y_{1i} \mid \text{ régimen "1" es observado}] &= E[Y_{1i} \mid W_i'\alpha > e_i] = \\ = X_i'\beta_1 + E[u_{1i} \mid W_i'\alpha > e_i] &= X_i'\beta_1 - \sigma_{1e}\, \phi_i/\Phi_i . \end{aligned} \tag{7}$$

Del mismo modo, si la probabilidad de estar en la muestra es muy grande, la expresión $\sigma_{1e}\phi_i/\Phi_i$ será pequeña, por lo que el problema de sesgo sería despreciable.

Supongamos que para el individuo i el régimen "1" es observado. Si el modelo probit pronostica mal para este individuo, entonces el valor de ϕ_i/Φ_i sería grande. En este caso, las expresiones $W_i'\alpha$ y e_i serían negativas. Si el parámetro σ_{1e} es negativo, entonces se esperaría que en promedio u_{1i} fuera positivo.

Para resolver el problema de sesgo por elección se pueden seguir dos procedimientos. El primero consiste en llevar a cabo la estimación por un método bietápico,[10] mientras que el segundo consiste en utilizar máxima verosimilitud.

El método bietápico a veces se utiliza como un procedimiento para obtener valores iniciales para los parámetros del programa de maximización de la verosimilitud. Sin embargo, el método bietápico también es utilizado como la etapa final en la estimación. Este método consiste en incorporar en la estimación de las ecuaciones de ingreso una estimación del valor esperado de los residuales, dado el proceso de selección. Definiendo

$$\begin{aligned} \Theta_{0i} &= \phi(W_i'\alpha)/[1 - \Phi\,(W_i'\alpha)] \\ \Theta_{1i} &= \phi(W_i'\alpha)/\Phi\,(W_i'\alpha), \end{aligned} \tag{8}$$

las ecuaciones de ingreso por estimar serían:

$$\begin{aligned} \ln(Y_{0i}) &= X_i'\beta_0 + \sigma_{0e}\Theta_{0i} + v_{0i} \\ \ln(Y_{1i}) &= X_i'\beta_1 - \sigma_{1e}\Theta_{1i} - v_{1i} \end{aligned} \tag{9}$$

donde,

$$\begin{aligned} u_{0i} &= \sigma_{0e}\Theta_{0i} + v_{0i}, \qquad & E[v_{0i} \mid W_i'\alpha \le e_i] &= 0 \\ u_{1i} &= -\,\sigma_{1e}\Theta_{1i} + v_{1i}, \qquad & E[v_{1i} \mid W_i'\alpha > e_i] &= 0. \end{aligned} \tag{10}$$

[9] De este modo, si existe una gran cantidad de individuos para los cuales el pronóstico del probit es malo y a la vez estos individuos tienen ingresos superiores a la media, entonces el estimador de σ_{0e} probablemente será positivo.

[10] Éste es el método de J. Heckman (1977).

De los primeros momentos de los residuales de (6) se tiene:[11]

$$
\begin{aligned}
E[u_{0i} \mid W_i'\alpha \leq e_i] &= \sigma_{0e}\Theta_{0i} \\
E[u_{1i} \mid W_i'\alpha > e_i] &= -\sigma_{1e}\Theta_{1i}
\end{aligned}
\tag{11}
$$

$$
\begin{aligned}
E[u_{0i}^2 \mid W_i'\alpha \leq e_i] &= \sigma_0^2 + \sigma_{0e}^2(W_i'\alpha)\Theta_{0i} \\
E[u_{1i}^2 \mid W_i'\alpha > e_i] &= \sigma_1^2 - \sigma_{1e}^2(W_i'\alpha)\Theta_{1i}
\end{aligned}
\tag{12}
$$

por lo que

$$
\begin{aligned}
\mathrm{Var}[v_{0i} \mid W_i'\alpha \leq e_i] &= \sigma_0^2 + \sigma_{0e}^2\Theta_{0i}(W_i'\alpha - \Theta_{0i}) \\
\mathrm{Var}[v_{1i} \mid W_i'\alpha > e_i] &= \sigma_1^2 - \sigma_{1e}^2\Theta_{1i}(W_i'\alpha + \Theta_{1i}),
\end{aligned}
\tag{13}
$$

lo cual indica que el sistema (9) es heterocedástico, por lo que el uso de mínimos cuadrados generalizados (MCG) incrementa la eficiencia de los estimadores bietápicos.

La estimación bajo este método bietápico consiste, primero, en estimar la ecuación (3) usando métodos probit. Una vez obtenidos estimadores de los parámetros α, obtener estimaciones para las expresiones en (8). Finalmente, incorporar estas últimas expresiones en las ecuaciones de ingresos (9) para después estimarlas por MCO o MCG.

En las ecuaciones (9) tiene especial importancia la estimación de los parámetros σ_{0e} y σ_{1e}, los cuales nos indican, cuando son estadísticamente diferentes de cero, la existencia de sesgo por elección en las estimaciones simples. Asimismo, el signo de estos parámetros nos indica si para cada régimen se está observando la cola inferior o superior de la distribución de ingresos.

Las varianzas de los parámetros de las ecuaciones de ingreso no corresponden a las reportadas por el método de MCO; sin embargo, los resultados numéricos son muy similares. La expresión para la matriz de varianzas y covarianzas de los parámetros en (9), la cual se utilizó en el presente trabajo, se encuentra en L. Lee, G. Maddala y R. Trost (1980).

Como se puede ver, este método bietápico es un método recursivo. Primero se estima la ecuación de elección, modelo probit, para, con los resultados de este modelo, estimar después las ecuaciones de ingreso. Sin embargo, el proceso termina en este punto, ya que no se vuelve a estimar la ecuación probit una vez estimadas las ecuaciones de ingreso. En el segundo método, el cual es utilizado en el presente trabajo, la estimación se lleva a cabo de manera conjunta.

Este último método consiste en encontrar el valor de los parámetros de interés, es decir, de los parámetros de las ecuaciones (1), (2) y (3), que maximizan la función de verosimilitud para este problema.

Al igual que en un caso bietápico, se supone que los residuales tienen una función

[11] Véase G. Maddala (1983) o L. Lee y R. Trost (1978).

normal trivariada. En este caso la función de verosimilitud se puede escribir del siguiente modo:[12]

$$L = \prod_{i=1}^{n} \left[\text{Prob}\,(u_{0i} \mid e_i \geq W_i'\alpha)\ \text{Prob}\,(e_i \geq W_i'\alpha)\right]^{1-I_i} \times \left[\text{Prob}\,(u_{1i} \mid e_i < W_i'\alpha)\ \text{Prob}\,(e_i < W_i'\alpha)\right]^{I_i}. \tag{14}$$

Sustituyendo las expresiones para estas probabilidades se llega a[13]

$$L = \prod_{i=1}^{n} \left[\int_{W_i'\alpha}^{\infty} \mathrm{f}\,(y_{0i} - X_i'\beta,\, e_i)\, \mathrm{d}e_i\right]^{1-I_i} \times \left[\int_{-\infty}^{W_i'\alpha} \mathrm{f}\,(y_{1i} - X_i'\beta,\, e_i)\, \mathrm{d}e_i\right]^{I_i} \tag{15}$$

en la cual f (u_{ji}, e_i) representa la función de densidad normal bivariada. A esta función hay que incorporar las restricciones de que se habló anteriormente, es decir, esperanzas incondicionales igual a cero, y varianza de e_i igual a la unidad. Utilizando resultados conocidos sobre distribuciones normales condicionadas, el logaritmo natural de esta función se puede escribir de la siguiente manera:

$$\log(L) = \sum_{i=1}^{n} \left\{ \begin{array}{l} I_i \log\left[\phi\left(\dfrac{y_{1i} - x_i'\beta_1}{\sigma_1}\right)\Phi\,(m_{1i})/\sigma_1\right] + \\ (1 - I_i)\log\left[\phi\left(\dfrac{y_{0i} - x_i'\beta_0}{\sigma_0}\right)(1 - \Phi\,(m_{0i}))/\sigma_0\right] \end{array} \right\} \tag{16}$$

donde,

$$m_{1i} = \frac{W_i'\alpha - (y_{1i} - x_1'\beta_1)\,\sigma_{1e}/\sigma_1^2}{\sqrt{1 - \sigma_{1e}^2/\sigma_1^2}} \tag{17}$$

$$m_{0i} = \frac{W_i'\alpha - (y_{0i} - x_1'\beta_0)\,\sigma_{0e}/\sigma_0^2}{\sqrt{1 - \sigma_{0e}^2/\sigma_0^2}}. \tag{18}$$

Al igual que en el método bietápico, para efectos de evaluar el sesgo por elección, tienen especial importancia los estimadores de las covarianzas de los residuales.

[12] Utilizando el hecho de que u_{0i} y u_{1i} no son observables simultáneamente.

[13] Véase, por ejemplo, G. Maddala (1983) o L. Lee (1979).

Resultados

Para llevar a cabo la estimación se utilizó información original de la *Encuesta Nacional de Ingreso-Gasto de los Hogares 1989* (ENIGH89). En esta encuesta existe información sobre diferentes características de los individuos, tales como nivel educativo, sueldo, edad, zona de residencia, etc., información que es muy importante para la estimación de las ecuaciones de ingreso y de elección. Sin embargo, la ENIGH89 no incluye información para todos los individuos sobre ciertas características familiares.

Datos sobre el nivel de educación de los padres, nivel de ingreso de la familia, tamaño de la familia, etc., son muy importantes para explicar la elección de los hijos de continuar o no con estudios superiores. Por esta razón, fue necesario limitar la muestra por utilizar, de modo que se tuviera información sobre las características familiares de los individuos. Los acotamientos hechos a la muestra son los siguientes:[14] *a)* sólo se incluyeron individuos que estuvieran viviendo con sus padres al momento del levantamiento de la encuesta, *b)* solamente se incluyó a individuos que dijeron tener un sueldo positivo, *c)* la edad de los individuos se limitó al rango 23-49 años, *d)* finalmente, sólo se consideró a los individuos de los cuales se tuviese información sobre todas las variables por incorporar.

Con estas acotaciones se alcanzó una muestra de 445 individuos que se dividió en dos grandes grupos: individuos con educación máxima de preparatoria o equivalente (régimen "0") e individuos con estudios superiores completos o incompletos (régimen "1"). El primer grupo lo integraron 214 individuos (48.1%) y el segundo 231 (51.9%).

Las variables utilizadas para modelar la elección de escolaridad son comunes en este tipo de estudios:[15]

1) COLLG: variable dicotómica que toma el valor unitario cuando el individuo elige educación superior, y cero en caso contrario. Ésta es la variable dependiente.

2) TAMH: número de miembros en la familia.

3) ESCOLJ: número de años de educación formal del padre.

4) TECJ: variable dicotómica que toma el valor unitario cuando el padre ha tenido algún tipo de preparación técnica.

5) LWJ: logaritmo del ingreso del padre.

6) ESCOLE: número de años de educación formal de la madre.

7) TECE: variable dicotómica que toma el valor unitario cuando la madre ha tenido algún tipo de preparación técnica.

8) SEXO: variable dicotómica que es igual a la unidad cuando es hombre, y cero si es mujer.

9) URBBAJO: variable dicotómica que toma el valor unitario cuando el individuo vive en zona urbana pero no es URBALTO.

10) URBALTO: variable dicotómica que toma el valor unitario cuando el individuo vive en el área metropolitana de la ciudad de México, Guadalajara o Monterrey.

[14] Un acotamiento similar se llevó a cabo para la estimación presente en A. Zamudio y T. Bracho (1994).

[15] Para una mayor discusión sobre cómo se crearon estas variables, véase T. Bracho y A. Zamudio (1994b).

Cuadro 1
Estadísticas sobre las variables utilizadas

	Total		*Educación media*		*Educación superior*	
	(a)	*(b)*	*(a)*	*(b)*	*(a)*	*(b)*
TAMH	6.78	2.30	7.30	2.45	6.31	2.05
ESCOLJ	6.04	3.92	5.01	2.82	7.00	4.52
LWJ	13.19	1.02	13.03	0.85	13.34	1.13
TECJ	0.09	0.28	0.07	0.25	0.11	0.31
ESCOLE	5.22	3.71	4.24	3.20	6.13	3.91
TECE	0.14	0.35	0.11	0.32	0.17	0.38
GÉNERO	0.49	0.50	0.37	0.48	0.61	0.49
LW	13.10	0.59	12.89	0.52	13.28	0.59
URBBAJO	0.57	0.50	0.54	0.50	0.60	0.49
URBALTO	0.36	0.48	0.37	0.48	0.35	0.48
EX	7.24	4.18	8.71	4.28	5.89	3.60
LHT	3.68	0.33	3.67	0.32	3.68	0.34

NOTA: (a) Promedio.
(b) Desviación Estándar.

Las variables explicativas de la ecuación de elección son consideradas como exógenas al modelo. Este supuesto puede presentar problemas puesto que las dos últimas variables, es decir, las que denotan la zona de residencia, pueden ser consideradas como endógenas, ya que el individuo puede elegir la zona de residencia.

Las variables utilizadas para las ecuaciones de ingreso, comunes en la estimación de ecuaciones del tipo "minceriana", son las siguientes:

1) LW: logaritmo natural del sueldo mensual del individuo. Es la variable dependiente.
2) EX: experiencia laboral.[16]
3) EX^2: el cuadrado de EX.
4) SEXO, URBBAJO y URBALTO: definidas anteriormente.
5) LHT: logaritmo natural de las horas trabajadas en una semana.

Las variables utilizadas tanto para la ecuación de elección como para las ecuaciones de ingreso son bastante comunes en este tipo de estudios. Variables como la educación de los padres, su ingreso, la zona de residencia y el género resultan importantes en la decisión de continuar con estudios superiores. De igual modo la experiencia laboral, las horas trabajadas, el género o la zona de residencia son variables muy utilizadas para explicar el nivel de ingresos.

[16] Como la ENIGH89 no contiene información sobre la experiencia laboral, se tuvo que construir dicha variable del modo usual en este tipo de trabajos, es decir, EX = edad – años de educación formal –6.

Cuadro 2
Ecuación de ingresos

	Educación media			*Educación superior*		
	Variable dependiente: LW					
Variables	*(a)*	*(b)*	*(c)*	*(a)*	*(b)*	*(c)*
CONSTANTE	11.3160**	11.1009**	11.1946**	11.4328**	12.0846**	12.4218**
	(0.4417)	(0.4354)	(0.2170)	(0.4669)	(0.4992)	(0.4052)
EX	0.0463	0.0388	0.0425	0.0211	0.0297	0.0267
	(0.0346)	(0.0339)	(0.0310)	(0.0321)	(0.0308)	(0.0267)
EX^2	–0.0012	–0.0010	–0.0011	–0.0003	–0.0007	–0.0004
	(0.0014)	(0.0014)	(0.0012)	(0.0018)	(0.0017)	(0.0015)
GÉNERO	0.0239	–0.1377	–0.0648	–0.0718	–0.2435*	–0.2808**
	(0.0725)	(0.0964)	(0.0705)	(0.0763)	(0.0982)	(0.0749)
URBBAJO	0.2880*	0.2042**	0.2480*	0.4752**	0.3136+	0.2486
	(0.1252)	(0.1415)	(0.1179)	(0.1709)	(0.1836)	(0.1622)
URBALTO	0.3367**	0.3038*	0.3286**	0.5918**	0.4476*	0.3784*
	(0.1294)	(0.1433)	(0.1226)	(0.1759)	(0.1870)	(0.1677)
LHT	0.2722*	0.2857**	0.2756**	0.3515**	0.3202**	0.2717**
	(0.1108)	(0.1058)	(0.0705)	(0.1095)	(0.1061)	(0.0909)
σ_{ie}	—	0.4581**	0.2600**	—	0.4801**	0.5834**
		(0.1608)	(0.0744)		(0.1558)	(0.0750)
$n =$	214	214		231	231	
$R^2 =$	0.085	0.124		0.093	0.136	
$F =$	3.187	4.148		3.833	5.033	

NOTA: + Significativo al 10%.
* Significativo al 5%.
** Significativo al 1%.
(a) Mínimos cuadrados ordinarios.
(b) Heckit.
(c) Máxima verosimilitud.
Error estándar entre paréntesis.

En el cuadro 1 se muestran algunas estadísticas descriptivas de dichas variables. Las estadísticas se presentan tanto para el total de la muestra como para las submuestras correspondientes a los individuos con educación media y superior.

En el cuadro 1 se puede ver que, en términos generales, los individuos que continuaron con estudios superiores provienen de familias menos numerosas, donde es mayor el promedio de educación de los progenitores, el ingreso del padre así como el porcentaje de padres que tienen algún tipo de preparación técnica. En cuanto a las características de los individuos, se observa que un mayor porcentaje de hombres con-

Cuadro 3
Ecuación de elección

Variables	*Modelo probit* Coeficiente	*Modelo probit* Error est.	*Estimación conjunta* Coeficiente	*Estimación conjunta* Error est.
CONSTANTE	–0.4648	0.9345	–2.7008**	0.6821
TAMH	–0.0964**	0.0287	–0.0840**	0.0220
ESCOLJ	0.0507*	0.0230	0.0235	0.0178
TECJ	0.0588	0.2491	0.0230	0.2171
LWJ	0.0161	0.0721	0.1894**	0.0516
ESCOLE	0.0576*	0.0252	0.0583**	0.0185
TECE	–0.2835	0.2152	–0.2292	0.1808
GÉNERO	0.6288**	0.1266	0.5960**	0.1049
URBBAJO	0.1659	0.2516	0.2063	0.2289
URBALTO	–0.0091	0.2619	0.0415	0.2393

NOTA: + Significativo al 10%.
* Significativo al 5%.
** Significativo al 1%.

tinuaron con estudios superiores, mientras que la zona de residencia no explica claramente el logro educativo. Respecto al sueldo, se ve claramente que los individuos con educación superior ganan, en promedio, más.

La estimación de las ecuaciones de ingreso se presenta en el cuadro 2. Con el objeto de comparar los diferentes métodos y ver claramente el problema de sesgo, las ecuaciones de ingreso se estimaron de acuerdo con tres diferentes métodos: el de los mínimos cuadrados ordinarios (MCO), el bietápico de Heckman o "Heckit"[17] y el de máxima verosimilitud (MV).

La estimación por MCO es la más sencilla de implementar: no se requiere ninguna estimación previa ni tampoco de estimación conjunta de las ecuaciones de ingreso y elección. Los resultados se reportan en la primera y cuarta columnas del cuadro 2.

Para utilizar el método "Heckit" se requiere, como se explicó en la sección anterior, estimar primeramente la ecuación de elección, es decir, el modelo probit. Los resultados de dicho modelo se presentan en el cuadro 3, primera y segunda columnas. Con esos resultados se estimaron las ecuaciones de ingresos corregidas. Los resultados de esta última etapa son los reportados en la segunda y quinta columnas del cuadro 2.

La estimación por MV se lleva a cabo de manera simultánea, esto es, tanto la ecuación de elección como las de ingreso se estiman conjuntamente. Para el programa de maximización se utilizaron como valores iniciales los resultados del procedimiento "Heckit". Los resultados de la estimación se presentan separadamente. Los parámetros

[17] Los resultados obtenidos por el segundo método son, además, utilizados como valores iniciales para la estimación hecha según el tercer método.

de la ecuación de elección se indican en la tercera y cuarta columnas del cuadro 3, mientras que los resultados de las ecuaciones de ingreso se presentan en la tercera y sexta columnas del cuadro 2.

La estimación de la ecuación de ingreso para los individuos con solamente educación media indica que, en general, los tres métodos utilizados arrojan resultados similares. Los signos de las variables explicativas son los esperados, y la precisión de los estimadores es bastante buena en algunos casos.

Los coeficientes de las variables EX y EX^2 no resultaron significativos, aunque la combinación de signos positivo y negativo, respectivamente, es la esperada. Dicha combinación es consistente con la trayectoria de ingresos cóncava, esto es, creciente durante los primeros años después de terminar los estudios, y decreciente al finalizar la vida laboral del individuo. En relación con estas variables, los tres métodos arrojan resultados muy similares. La explicación de que los estimadores de estos coeficientes no resultaran significativos se encuentra en la muestra utilizada para la estimación. Puesto que estamos utilizando exclusivamente individuos que viven con sus padres, no existe información suficiente sobre individuos mayores de 30 años, por lo que no es posible estimar con precisión las trayectorias cóncavas de los ingresos.

La variable que denota al género (SEXO) resultó poco significativa en los tres métodos. De hecho, el signo del coeficiente de esta variable cambia cuando se ajusta por sesgo, es decir, cuando se utilizan los métodos "Heckit" y MV. Una explicación reside en la poca precisión que se tiene para el estimador de este coefiente. Sin embargo, también es posible pensar que los ingresos de la mujer, una vez hecho el ajuste por la probabilidad de acceder al mercado laboral, son al menos iguales a los del hombre.[18]

Los coeficientes de las variables URBBAJO y URBALTO resultan significativos y con los signos esperados. Ambos coeficientes se comportan de manera similar en cualquiera de los tres métodos: el de URBALTO es mayor que el de URBBAJO. Ambos coeficientes son menores cuando se lleva a cabo el ajuste por sesgo, aunque esta disminución no es muy grande y se puede explicar por el mismo ajuste que se hace, es decir, se ajusta por la probabilidad de ser observado en ese régimen, ya que en parte esta probabilidad refleja la zona de origen del individuo.

La variable LHT resulta positiva y significativa, ya que la magnitud del coeficiente se incrementa ligeramente cuando se ajusta por sesgo.

Finalmente, el coeficiente de σ_{0e} resulta positivo y significativo; lo primero indica la existencia de sesgo por elección en la estimación simple de la ecuación de ingresos para estudios medios; lo segundo, que se está observando la cola superior de la distribución de ingresos para estudios medios.

La ecuación de ingresos para estudios superiores arroja resultados similares a los del caso anterior. Los estimadores de los coeficientes de las variables EX y EX^2 no son significativos, aunque resultan con la combinación de signos esperados.

El signo del estimador del coeficiente de la variable GÉNERO resulta negativa en los

[18] Obviamente, hay que tomar en cuenta las acotaciones que se hicieron a la muestra con tal de llevar a cabo la estimación, ya que no permiten generalizar los resultados a toda la población.

tres métodos, y significativa cuando se ajusta por sesgo. Una diferencia importante se da en relación con la variable URBBAJO, la cual deja de ser significativa según los métodos "Heckit" y MV.[19] Los estimadores de los coeficientes de URBALTO y LIIT resultan positivos y significativos según los tres métodos.

El estimador del coeficiente $\sigma_{1\varepsilon}$ resulta positivo y significativo. Al igual que el caso anterior, el de la educación media, el ser significativo revela la existencia de sesgo por elección. El hecho de ser positivo indica, en este caso, que se está observando la cola inferior de la distribución de ingresos para estudios superiores.

La ecuación de elección se estimó con dos métodos. En el primero se estimó un modelo probit, que fue la base para el ajuste por sesgo que se hizo en las ecuaciones de ingreso. El segundo consistió en una estimación conjunta tanto de la ecuación de elección como de las de ingreso.

Para esta ecuación de elección algunas de las variables explicativas resultaron no significativas, aunque en general el signo fue el esperado.

El coeficiente de TAMH resultó negativo y significativo, como se esperaba. Esto es, a mayor tamaño de familia menor probabilidad de los hijos de ir más allá de los estudios medios. Cabe hacer notar que el coeficiente resultó menor en términos absolutos en la estimación conjunta.

Las variables ESCOLJ y ESCOLE resultaron positivas y significativas. LWJ resultó positiva pero no significativa en el modelo probit, sin embargo, en la estimación conjunta, incrementó su magnitud y se hizo significativa. Este cambio en la magnitud del coeficiente de LWJ se dio junto con una disminución del coefiente del ESCOLJ, el cual dejó de ser significativo. Esto se puede explicar porque ambas variables se encuentran muy correlacionadas, por lo que, en cierto modo, explican lo mismo.

La variable SEXO resultó, como se esperaba, positiva y significativa; las variables TECJ y TECE, no significativas, lo que indica que la preparación técnica del padre o madre no incrementa la probabilidad de continuar con estudios superiores.

Se esperaría que la zona de residencia fuera una variable importante en la probabilidad de continuar con estudios superiores. Sin embargo, las variables URBBAJO y URBALTO resultaron positivas pero no significativas. Esta situación se explica por la manera como se definieron las variables de zona; tal vez debió utilizarse otra.[20]

A partir de los resultados del método de máxima verosimilitud, es decir, la estimación conjunta, se llevaron a cabo algunos ensayos de hipótesis sobre igualdad de parámetros. Para realizarlos se utilizó la prueba de razón de verosimilitud.

Primeramente se ensayó sobre la igualdad, entre las dos ecuaciones de ingresos, de los parámetros de las variables de experiencia laboral (EX y EX^2). Esta igualdad es importante ya que, de ser cierta, facilita mucho el cálculo de los retornos.

Después se ensayó sobre la igualdad entre todos los parámetros de las dos ecuaciones de ingreso, con excepción del término constante y de la covarianza entre residuales.

[19] Una explicación posible se dio cuando se discutió la ecuación de ingresos para estudios medios.

[20] También hay que tomar en cuenta que las variables de zona indican el punto de residencia en el momento de la entrevista, lo que no necesariamente coincide con la zona de residencia en el momento de tomar la decisión sobre continuar o no con su educación.

Cuadro 4
Tasa de retorno a la educación superior

	Total (%)	*Hombres (%)*	*Mujeres (%)*
Total	26.93	24.10	29.67
Rural	26.46	23.62	29.20
Urbano bajo	26.48	23.64	29.21
Urbano alto	27.75	24.92	30.48

En ambos casos no fue posible rechazar la hipótesis sobre igualdad de los parámetros de las ecuaciones de ingreso.[21] De este modo los retornos a la educación superior pueden ser aproximados de una manera fácil. Sin embargo, en este trabajo los retornos son calculados usando integración numérica.

Con los resultados de la estimación se calcularon tasas de retorno para la educación superior y, en dicha operación, no se tomaron en cuenta los costos directos de la educación ni los ingresos que los alumnos de educación superior pudieran tener al estar estudiando. De esta manera se espera que ambos conceptos se cancelen de manera aproximada.[22]

Las tasas de retorno se obtuvieron al resolver numéricamente para r la siguiente ecuación:

$$V_0 = \int_0^n Y_0 (X_i, t)\, e^{-rt} dt = \int_0^{n-s} Y_1 (X_i, t)\, e^{-r(t+s)}\, dt = V_1 .$$

Donde n, el año de retiro de los individuos (65 años), y s, el tiempo que toma la educación superior (4 años). Las funciones de ingresos para ambos regímenes se calcularon con los parámetros estimados anteriormente.

En este caso las tasas de retorno van a quedar en función de los valores que tomen las variables explicativas diferentes a la experiencia laboral. Por esta razón se tomaron escenarios distintos para estas variables.

Utilizando los resultados del cuadro 2 y tomando los promedios muestrales de las variables explicativas GÉNERO, URBBAJO, URBALTO y LHT se obtuvo un retorno de 10.45%[23] cuando se hizo el cálculo con los estimadores de MCO, es decir, sin ajustar por el sesgo por elección. El retorno fue de 28.62% cuando se usó "Heckit", y de 26.93%

[21] Para el primer ensayo, el estadístico de la razón de verosimilitud es 0.13, el cual tiene teóricamente una distribución chi-cuadrada con 2 grados de libertad. Para el segundo ensayo, el estadístico toma el valor de 4.19, que tiene teóricamente una distribución chi-cuadrada con 6 grados de libertad.

[22] Esta forma de abordar el problema de los costos directos no es nueva. Véase, por ejemplo, J. Mincer (1974).

[23] En T. Bracho y A. Zamudio (1994b) se estimaron retornos para todo el ciclo escolar y para una muestra menos restringida que la del presente trabajo; los resultados arrojaron un retorno de 14.88 y 14.06% para la educación superior incompleta y completa, respectivamente.

con MV. Al comparar estos resultados se ven los efectos que tiene el sesgo por elección en la estimación de los retornos para la muestra considerada.

En el cuadro 4 se presentan, de acuerdo con diferentes escenarios de las variables explicativas los retornos calculados; con base en los estimadores de máxima verosimilitud.

En él se puede ver que en términos generales los retornos resultaron altos, siendo mayores para las mujeres y para los habitantes de las tres principales zonas metropolitanas de México.

Conclusiones

En este trabajo se llevó a cabo una estimación de los retornos a la educación superior en México libres del problema de sesgo por elección.

Los resultados indican que el fenómeno de la endogeneidad de la educación, o problema de selectividad, es muy importante tanto para la educación superior como para la media.

Este resultado es congruente con otros estudios. En el estudio de A. Zamudio y T. Bracho (1994) se encontró que el problema de la selectividad era importante en una estimación llevada a cabo para todo el ciclo escolar. L. Kenny, L. Lee, G. Maddala y R. Trost (1979) no encontraron problema de selectividad para la educación superior, pero sí para la educación media. R. Willis y S. Rosen (1979) encontraron problemas de selectividad tanto para la educación superior como para la media. E. Cohn y W. Hughes (1994) encontraron problemas de selectividad tanto para la educación media como para la superior, y también obtuvieron retornos significativamente superiores cuando ajustaron las ecuaciones de ingreso por sesgo por elección.

Al calcular los retornos se vio la diferencia que hace el ajustar o no las ecuaciones de ingreso por selectividad: los retornos casi se triplican cuando se ajustan por sesgo.

Este resultado, aunque importante, debe ser tomado con reservas. El principal problema radica en la muestra utilizada para la estimación, ya que contiene muchas restricciones, lo cual no permite que los resultados puedan ser generalizados a toda la población.[24] También hay que tomar en cuenta los posibles problemas de heterocedasticidad en la muestra. Parte de estos problemas se trataron al ajustar las ecuaciones de ingreso por selectividad, sin embargo, es posible que existan otras fuentes de heterogeneidad en la muestra.

[24] Hay que recordar que la muestra utilizada para la estimación estuvo formada por individuos que todavía vivían con sus padres.

Referencias bibliográficas

Bracho, Teresa y Andrés Zamudio (1994), *Rendimientos económicos a la escolaridad I: discusión teórica y métodos de estimación*, México, CIDE (Documento de Trabajo 30, E).

——— (1994), *Rendimientos económicos a la escolaridad II: estimaciones para el caso mexicano, 1989*, México, CIDE (Documento de Trabajo 31, E).

Carnoy, M. (1967), "Rates of Return to Schooling in Latin America", *Journal of Human Resources*, vol. 2, núm. 3, pp. 359-374.

Cohn, Elchanan y Woodrow Hughes (1994), "A Benefit-Cost Analisys of Investment in College Education in the United States: 1969-1985", *Economics of Education Review*, vol. 13, núm. 2, pp. 109-123.

Griliches, Z. (1977), "Estimating the Returns to Schooling: Some Econometric Problems", *Econometrica*, vol. 45, núm. 1, pp. 1-22.

Heckman, James (1979), "Sample Selection Bias as a Specification Error", *Econometrica*, vol. 47, núm. 1, pp. 153-161.

Kenny, Lawrence, Lung-Fei Lee, G. Maddala y R. Trost (1979), "Returns to College Education: An Investigation of Self Selection Bias Based on the Project Talent Data", *International Economic Review*, vol. 20, núm. 3, pp. 775-789.

Lee, Lung-Fei (1978), "Unionism and Wage Rates: A Simultaneous Equations Model with Qualitative and Limited Dependent Variables", *International Economic Review*, vol. 19, núm. 2, pp. 415-433.

——— (1979), "Identification and Estimation in Binary Choice Models with Limited (Censored) Dependent Variables", *Econometrica*, vol. 47, núm. 4, pp. 977-996.

Lee, Lung-Fei, G. Maddala y R. Trost (1980), "Asymtotic Covariance Matrices of Two-Stage Probit and Two-Stage Tobit Methods for Simultaneous Equations Models with Selectivity", *Econometrica*, vol. 48, núm. 2, pp. 491-503.

Lee, Lung-Fei y Robert P. Trost (1978), "Estimation of Some Limited Dependent Variable Models with Application to Housing Demand", *Journal of Economterics*, núm. 3, pp. 357-382.

Maddala, G. (1983), *Limited-Dependent and Qualitative Variables in Econometric*, Cambridge University Press.

Mincer, J. (1974), *Schooling, Experience and Earnings*, National Bureau of Economic Research.

Psacharopoulos, G. y Y. Chu Ng (1992), *Earnings and Education in Latin America: Assessing Priorities for Schooling Investments*, Banco Mundial (Policy Research Working Paper, WPS 1056).

Willis, Robert (1986), "Wage Determinants: A Survey and Reinterpretation of Human Capital Earnings Functions", Ashenfelter O. y R. Layard, *Handbook of Labor Economics*, vol. 1, Elsevier Science.

Willis, Robert y Sherwin Rosen (1979), "Education and Self Selection", *Journal of Political Economy*, vol. 87, núm. 5, pp. s7-s36.

Zamudio, Andrés y Teresa Bracho (1994), *Rendimientos económicos a la escolaridad III: el problema de sesgo por elección*, México, CIDE (Documento de Trabajo 31, E).