

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

Andrés Zamudio Carrillo

Resumen: En el presente artículo se estiman las tasas de retorno a la educación superior en México tomando en cuenta el posible sesgo por elección del individuo. El método utilizado es el de máxima verosimilitud. La información utilizada consiste en una muestra de la *Encuesta Nacional de Ingresos y Gastos de los Hogares 1989*, la cual sólo incluye a individuos que aún viven con sus padres. Los resultados indican que el sesgo por elección resulta ser muy importante para la muestra utilizada, dando por resultado tasas de retorno significativamente más altas cuando se ajusta para corregir por este tipo de sesgo.

Abstract: This paper estimates the rates of return to college education in Mexico adjusting for self selection bias using maximum likelihood. Information is taken from the 1989 National Household Survey, where only individuals for whom information on parents is available are included in the sample. The results indicate the importance of the self selection bias for the sample used, giving rates of return significantly greater when the correction for this type of bias is applied.

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 acerca de los rendimientos a la educación en México. Como ejemplos de estos trabajos se encuentran los de

Profesor-investigador de la División de Economía del CIDE.

M. Carnoy (1967), G. Psacharopoulos e Y. Chu Ng (1992), T. Bracho y A. Zamudio (1994), A. Zamudio y T. Bracho (1994). En todos ellos, a diferencia del último, 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 el estudio de A. Zamudio y T. Bracho (1994) se estiman retornos a la educación libres del problema de sesgo por elección, sin embargo estos retornos se calculan para la educación en general, sin considerar las particularidades que puede tener cada ciclo educativo.

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 características de los individuos como factores socioeconómicos. Muchos de estos factores forman parte, a la vez, de la decisión de continuar con estudios superiores y del desempeño laboral del individuo. Puesto 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 dichas ecuaciones, 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.

Método de estimación³

Para llevar a cabo el cálculo de los rendimientos se estiman primero ecuaciones de ingreso para individuos con educación media (preparatoria) y para cuando se 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, la educación media, y con "1" a la segunda, la de 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, 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 estará denotada por $Y(S, X_i)$, donde S sólo asume 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 así:

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

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

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, se observa la ecuación (2).

¹ Una discusión acerca del problema de sesgo por elección se encuentra en A. Zamudio y T. Bracho (1994).

² Véase J. Heckman (1979).

³ La presente discusión respecto al método de estimación es conocida en la literatura sobre el tema, 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".

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

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

La educación puede ser vista tanto como una inversión como un consumo. De esta manera, variables como la educación de los padres o la zona de residencia pueden ser variables 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 por la educación.⁶

Estos factores consisten en variables observables y no observables o medibles. Denotemos los observables con el vector W_i y a los no observables por la variable aleatoria e_i , y definamos la variable latente I_i^* del siguiente modo:

$$I_i^* = W_i' \alpha - e_i. \quad (3)$$

La variable I_i^* no es observable en forma continua, sólo lo es 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,

$$I_i = 1 \quad \text{cuando} \quad I_i^* > 0$$

$$I_i = 0 \quad \text{de otro modo.} \quad (4)$$

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.

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

⁶ En algunos estudios como en R. Willis y S. Rosen (1979) o en 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 utilizarse información de corte transversal, se está suponiendo que los ingresos actuales son equivalentes a los ingresos esperados en el momento de tomar la decisión. Por supuesto, éste no sería el caso cuando se tiene una situación de cambio. Precisamente 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), sino que la elección se hizo en función de variables exógenas.

Como (2) se observa 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) se observa cuando $W_i' \alpha > e_i$. Del mismo modo, (1) se observa 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,

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

Para esta matriz de covarianzas tenemos las siguientes restricciones. Ya 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). Como la tercera ecuación es del tipo probit, la varianza para e_i no es identificable, por lo que 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 del valor del argumento $W_i' \alpha$, mientras que la probabilidad de observar el régimen "0" estará en función inversa del argumento.

Utilizando resultados conocidos para distribuciones normales truncadas⁷ la posibilidad de sesgo por elección para el régimen "0" se puede ver en las siguientes expresiones:

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

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

ϕ 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, pues 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.

Se puede demostrar que la expresión $\phi_i / (1 - \Phi_i)$ es una función creciente del argumento $W_i' \alpha$,⁸ la cual a su vez está en función indirecta 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 ese régimen.

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

⁸ Véase L. Lee (1979).

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", el estimador de σ_{0e} tenderá a ser positivo.

$$E[Y_{1i} | \text{régimen "1" es observado}] = E[Y_{1i} | W_i' \alpha > e_i] = X_i' \beta_1 + E[u_{1i} | W_i' \alpha > e_i] = X_i' \beta_1 - \sigma_{1e} \phi_i / \Phi_i. \quad (7)$$

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,

$$\Theta_{0i} = \phi(W_i' \alpha) / [1 - \Phi(W_i' \alpha)]$$

$$\Theta_{1i} = \phi(W_i' \alpha) / \Phi(W_i' \alpha) \quad (8)$$

las ecuaciones de ingreso por estimar serían

$$\ln(Y_{0i}) = X_i' \beta_0 + \sigma_{0e} \Theta_{0i} + u_{0i}$$

$$\ln(Y_{1i}) = X_i' \beta_1 - \sigma_{1e} \Theta_{1i} + u_{1i}, \quad (9)$$

¹⁰ Éste es el método de J. Heckman (1979).

donde

$$\begin{aligned} u_{0i} &= \sigma_{0e} \Theta_{0i} + v_{0i}, & E[v_{0i} | W_i' \alpha \leq e_i] &= 0 \\ u_{1i} &= -\sigma_{1e} \Theta_{1i} + v_{1i}, & E[v_{1i} | W_i' \alpha > e_i] &= 0. \end{aligned} \quad (10)$$

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

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

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

por lo que

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

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

La estimación en este método bietápico consiste en estimar, primero, la ecuación (3) usando métodos probit. Una vez obtenidos los estimadores de los parámetros α , obtener estimaciones para las expresiones en (8). Finalmente, incorporar estas últimas expresiones a 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 in-

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

dica si se está observando, para cada régimen, la cola inferior o superior de la distribución de ingresos que corresponde a cada régimen.

Las varianzas de los parámetros de las ecuaciones de ingreso no corresponden a las varianzas 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), que se utilizó en el presente trabajo, se puede ver en L. Lee, G. Maddala y R. Trost (1980).

Como puede verse, 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, utilizado en el presente trabajo, la estimación se lleva a cabo de manera conjunta.

Este segundo 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 el caso bietápico, se supone que los residuales tienen una función normal trivariada. En este caso la función de verosimilitud se puede escribir del siguiente modo:¹²

$$\begin{aligned} L &= \prod_{i=1}^n \left[\text{Prob}(u_{0i} | e_i \geq W_i' \alpha) \text{Prob}(e_i \geq W_i' \alpha) \right]^{1-I_i} \\ &\quad \times \left[\text{Prob}(u_{1i} | e_i < W_i' \alpha) \text{Prob}(e_i < W_i' \alpha) \right]^{I_i}. \end{aligned} \quad (14)$$

Sustituyendo las expresiones para estas probabilidades se llega a¹³

$$L = \prod_{i=1}^n \left[\int_{W_i' \alpha}^{\infty} f(y_{0i} - X_i' \beta_0, e_i) de_i \right]^{1-I_i} \times \left[\int_{-\infty}^{W_i' \alpha} f(y_{1i} - X_i' \beta_1, e_i) de_i \right]^{I_i}, \quad (15)$$

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

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).

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:

$$\log(L) = \sum_{i=1}^n \left\{ I_i \log \left[\phi \left(\frac{y_{1i} - X_i' \beta_1}{\sigma_1} \right) \Phi(m_{1i}) / \sigma_1 \right] + (1 - I_i) \log \left[\phi \left(\frac{y_{0i} - X_i' \beta_0}{\sigma_0} \right) (1 - \Phi(m_{0i})) / \sigma_0 \right] \right\}, \quad (16)$$

donde

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

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

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

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 acotar 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 incluyó a individuos que estuvieran viviendo con sus padres en el momento del levantamiento de la encuesta; b) sólo se incluyó a individuos que dijeran tener un sueldo positivo; c) la edad de los individuos se acotó al rango de 23 a 49 años, d) finalmente, sólo se consideraron los individuos para los cuales se tuviera información sobre todas las variables por incorporar.

Con estas acotaciones se llegó a una muestra de 445 individuos. Ésta 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 constó de 214 individuos (48.1%) y el segundo de 231 (51.9%).

Las variables utilizadas para modelar la elección por la escolaridad son variables 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 el 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 el entrevistado 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 discusión más amplia sobre cómo se crearon estas variables véase T. Bracho y A. Zamudio (1994b).

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 variables 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, y que son comunes a las de la estimación de ecuaciones del tipo "minceriana", son las siguientes:

- 1) LW: es el logaritmo natural del sueldo mensual del individuo. Ésta es la variable dependiente.
- 2) EX: experiencia laboral.¹⁶
- 3) EX2: el cuadrado de EX.
- 4) SEXO, URBAJO y URBALTO: las cuales se definieron 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 se consideran 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 comúnmente utilizadas para explicar el nivel de ingresos.

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

En este cuadro se puede ver que, en términos generales, los individuos que continuaron con estudios superiores provienen de familias menos numerosas, en las que el promedio de educación de los padres es mayor, el ingreso del padre es mayor y un mayor porcentaje de los padres tiene algún tipo de preparación técnica. En cuanto a las características de los individuos, se puede observar que un mayor porcentaje de hombres continuó con estudios superiores, mientras que la zona de residencia no explica claramente el logro educativo. En cuanto al suel-

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

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
URBAJO	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

Notas: (a) Promedio.
(b) Desviación estándar.

do, 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. A fin de comparar los diferentes métodos y ver claramente el problema de sesgo, las ecuaciones de ingreso se estimaron bajo tres diferentes métodos: mínimos cuadrados ordinarios (MCO), el método bietápico de Heckman o "Heckit"¹⁷ y por máxima verosimilitud (MV).

La estimación por MCO es la más sencilla de implementar. Para esta estimación no se requiere ninguna estimación previa, ni tampoco se llevó a cabo ninguna estimación conjunta de las ecuaciones de ingreso y elección. Los resultados se muestran en la primera y cuarta columnas del cuadro 2.

¹⁷ Los resultados producidos por el segundo método son, además, utilizados como valores iniciales para la estimación en el tercer método.

Cuadro 2. Ecuación de ingresos

Variables	Variable dependiente: LW					
	Educación media			Educación superior		
	(a)	(b)	(c)	(a)	(b)	(c)
Constante	11.3160** (0.4417)	11.1009** (0.4354)	11.1946** (0.2170)	11.4328** (0.4669)	12.0846** (0.4992)	12.4218** (0.4052)
EX	0.0463 (0.0346)	0.0388 (0.0339)	0.0425 (0.0310)	0.0211 (0.0321)	0.0297 (0.0308)	0.0267 (0.0267)
EX2	-0.0012 (0.0014)	-0.0010 (0.0014)	-0.0011 (0.0012)	-0.0003 (0.0018)	-0.0007 (0.0017)	-0.0004 (0.0015)
GÉNERO	0.0239 (0.0725)	-0.1377 (0.0964)	-0.0648 (0.0705)	-0.0718 (0.0763)	-0.2435* (0.0982)	-0.2808** (0.0749)
URBBAJO	0.2880* (0.1252)	0.2042** (0.1415)	0.2480* (0.1179)	0.4752** (0.1709)	0.3136+ (0.1836)	0.2486 (0.1622)
URBALTO	0.3367** (0.1294)	0.3038* (0.1433)	0.3286** (0.1226)	0.5918** (0.1759)	0.4476* (0.1870)	0.3784* (0.1677)
LHT	0.2722* (0.1108)	0.2857** (0.1058)	0.2756** (0.0705)	0.3515** (0.1095)	0.3202** (0.1061)	0.2717** (0.0909)
σ_{ϵ}	-	0.4581** (0.1608)	0.2600** (0.0744)	-	0.4801** (0.1558)	0.5834** (0.0750)
n	214	214	231	231	231	231
R^2	0.085	0.124	0.093	0.093	0.136	0.136
F	3.187	4.148	3.833	3.833	5.033	5.033

Notas: + significativo al 10 por ciento.

* significativo al 5 por ciento.

** significativo al 1 por ciento.

(a) Mínimos cuadrados ordinarios.

(b) Heckit.

(c) Máxima verosimilitud.

Error estándar entre paréntesis.

Para realizar la estimación por el método "Heckit" se requiere, como se explicó en la sección anterior, estimar, en primer lugar, la ecuación de elección, es decir, el modelo probit. Los resultados del modelo probit se presentan en el cuadro 3, primera y segunda columnas. Con estos resultados se estimaron las ecuaciones de ingresos corregidas. Los resultados de esta última etapa son los que se muestran en la segunda y quinta columnas del cuadro 2.

La estimación por MV se lleva a cabo de manera conjunta, esto es, tanto la ecuación de elección como las de ingreso se estiman al mismo tiempo. Para el programa de maximización se utilizaron como valores iniciales los resultados del procedimiento "Heckit". Los resultados de la estimación se presentan por separado. Los parámetros de la ecuación de elección se presentan en la tercera y cuarta columnas del cuadro 3, mientras que los resultados para 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 que únicamente tienen educación media indica lo siguiente. En general, los tres métodos utilizados arrojan resultados similares. Asimismo, los signos de las variables explicativas son los que se habían esperado, siendo que la precisión de los estimadores es bastante buena en algunos casos.

Los coeficientes de las variables EX y EX2 no resultaron significativos, aunque la combinación de signos positivo y negativo, respectivamente, es la esperada. Esta combinación es congruente 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 por la que los estimadores de estos coeficientes no resultaron significativos se encuentra en la muestra utilizada para la estimación. Puesto que estamos incluyendo sólo a individuos que viven con sus padres, no existe información suficiente sobre individuos relativamente mayores, digamos, mayores de 30 años, por lo cual 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 para este hecho reside en la poca precisión que se tiene para el estimador de este coeficiente. Sin embargo, también es posible pensar que los ingresos de la mujer, una

Cuadro 3. Ecuación de elección

Variables	Modelo probit		Estimación conjunta	
	Coefficiente	Error estándar	Coefficiente	Error estándar
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
URRBAJO	0.1659	0.2516	0.2063	0.2289
URBALTO	-0.0091	0.2619	0.0415	0.2393

Notas: + significativo al 10 por ciento.
* significativo al 5 por ciento.
** significativo al 1 por ciento.

vez ajustado por la probabilidad de incorporarse al mercado laboral, son al menos iguales a los del hombre.¹⁸

Los coeficientes de las variables URBAJO y URBALTO son significativos y con los signos esperados. Ambos coeficientes son muy similares en cualquiera de los tres métodos, siendo que el coeficiente de URBALTO es mayor que el coeficiente de URBAJO en cualquiera de los tres métodos. Ambos coeficientes son menores cuando se lleva a cabo el ajuste por sesgo, aunque esta disminución no es muy grande. Ésta puede explicarse por el mismo ajuste que se hace, se ajusta por la probabilidad de ser observado en ese régimen, siendo que, en parte, esta probabilidad refleja la zona de origen del individuo.

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

Finalmente, el coeficiente de σ_{0e} resulta positivo y significativo. El hecho de ser significativo indica la existencia de sesgo por elección en la estimación simple de la ecuación de ingresos para estudios medios. El hecho de ser positivo indica que se está observando a 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 EX2 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 tres métodos, siendo que ésta es significativa cuando se ajusta por sesgo. Una diferencia importante se da en relación con la variable URBAJO, la cual deja de ser significativa en los métodos "Heckit" y MV.¹⁹ Los estimadores de los coeficientes de URBALTO y LHT resultan positivos y significativos en los tres métodos.

El estimador del coeficiente σ_{1e} resulta positivo y significativo. Al igual que en 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.

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

¹⁹ Una explicación posible se proporcionó cuando se discutió la ecuación de ingresos para estudios medios.

La ecuación de elección se estimó mediante dos métodos. En el primero se estimó un modelo probit, el cual 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 tienen los hijos de ir más allá de los estudios medios. Es de 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 esta variable incrementó su magnitud y se hizo significativa. Este cambio en la magnitud del coeficiente de LWJ se dio en conjunto con una disminución del coeficiente del ESCOLJ, el cual dejó de ser significativo. Esto se puede explicar por el hecho de que las dos variables se encuentran muy correlacionadas, por lo que ambas explican, en cierto modo, lo mismo.

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

Se esperaba que la zona de residencia fuera una variable importante en la probabilidad de continuar con estudios superiores. Sin embargo, las variables URBAJO y URBALTO resultaron positivas, pero no significativas. Esta situación se puede explicar por la manera como se definieron las variables de zona, tal vez hubiera sido necesario utilizar otra definición de zona o región.²⁰

Utilizando los resultados del método de máxima verosimilitud, la estimación conjunta, se llevaron a cabo algunos ensayos de hipótesis sobre igualdad de parámetros. Para realizar los ensayos se utilizó la prueba de razón de verosimilitud.

Primero se ensayó sobre la igualdad, entre las dos ecuaciones de

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

ingresos, de los parámetros de las variables de experiencia laboral (EX y EX2). 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.

En ambos casos, no fue posible rechazar la hipótesis sobre igualdad de los parámetros de las ecuaciones de ingreso.²¹ De esta manera, 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.

Utilizando los resultados de la estimación se calcularon tasas de retorno para la educación superior. Para realizar el cálculo de estas tasas de retorno no se tomaron en cuenta los costos directos de la educación ni se consideraron los ingresos que los estudiantes de educación superior pudieran tener al estar estudiando. Al proceder 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 es la edad de retiro de los individuos, que fue de 65 años, y s el tiempo que toma la educación superior, que fue de 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 diferentes escenarios para estas variables.

Utilizando los resultados del cuadro 2 y tomando los promedios muestrales de las variables explicativas GÉNERO, URBBAJO, URBALTO

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

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

Cuadro 4. Tasa de retorno a la educación superior (porcentajes)

	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

y LHT, se obtuvo un retorno de 10.45%²³ cuando se hizo el cálculo con los estimadores de MCO, sin ajustar por el sesgo por elección. El retorno fue de 28.62% cuando se utilizó "Heckit", y de 26.93% con MV. Al comparar estos resultados, se pueden observar los efectos que tiene el sesgo por elección en la estimación de los retornos para la muestra considerada.

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

En este cuadro 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 A. Zamudio y T. Bracho (1994) se encontró que el problema de selectividad era importante, esto para una estimación llevada a cabo para todo el ciclo

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

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 ocasiona el ajustar o no las ecuaciones de ingreso por selectividad: los retornos casi se triplican cuando se ajusta por sesgo.

Este resultado, aunque importante, debe tomarse con reservas. El principal problema radica en la muestra utilizada para la estimación, ya que ésta contiene muchas restricciones, lo cual no permite que estos 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.

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 Analysis 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), "Asymptotic 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 Econometrics*, 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", en O. Ashenfelter y R. Layard, *Handbook of Labor Economics*, vol. I, Elsevier Science Publishers.
- 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).

²⁴ Hay que recordar que la muestra utilizada para la estimación consistió en individuos que todavía vivían con sus padres.