

CENTRO DE INVESTIGACIÓN Y DOCENCIA ECONÓMICAS, A.C.



LOS RETORNOS A LA EDUCACIÓN A NIVEL REGIONAL EN MÉXICO

T E S I N A

QUE PARA OBTENER EL TÍTULO DE:

MAESTRÍA EN ECONOMÍA

PRESENTA

MARITZA ROSALES REYES

ASESOR DE TESIS:

DR. ROBERT DUVAL HERNÁNDEZ

MÉXICO, D.F.

JUNIO 2008

A mis padres

*Se agradecen de manera especial los comentarios y la ayuda
de los Doctores Robert Duval y Ricardo Smith.*

ÍNDICE

Introducción

1. Revisión bibliográfica.....	3
1.1 Estudios para México.....	5
2. El modelo	12
3. Estadísticas descriptivas.....	16
4. Resultados del modelo lineal clásico	22
5. Resultados del modelo de variables instrumentales.....	30

Conclusiones

Anexo 1

Anexo 2

Anexo 3

Bibliografía

INTRODUCCIÓN

La interrupción del proceso de convergencia regional en México a mediados de la década de los ochenta y el inicio de uno de divergencia y polarización regional a partir de entonces (cuando el país se abre al comercio internacional), implica una enorme oportunidad para que el sector público intervenga a través de acciones que favorezcan la igualación de los niveles de desarrollo entre las distintas zonas. La literatura sobre crecimiento regional hace referencia de manera importante al papel de la inversión en infraestructura y en capital humano en el desarrollo económico local, y como herramienta para favorecer la convergencia.¹ No obstante, como en casi todos los problemas económicos, la escasez de los recursos enfrenta la disyuntiva entre dar prioridad a la eficiencia o a la equidad en las políticas públicas. En este caso, muchas de las acciones que favorecen la convergencia regional pueden implicar una contradicción con el objetivo de eficiencia.

En particular, atendiendo el debate sobre rendimientos crecientes o decrecientes de la educación, es necesario verificar si a escala regional las zonas de mayor nivel educativo (y eventualmente de ingreso promedio) muestran también mayores tasas de retorno de la escolaridad, o si son las regiones con los niveles educativos más bajos las que presentan los mayores rendimientos. En el caso de que la rentabilidad de la educación sea más elevada en las regiones de menor nivel educativo promedio, los objetivos de equidad y eficiencia no entran en conflicto. Esto se debe a que al destinar preferentemente las inversiones en educación hacia las regiones de menor desarrollo se atendería el objetivo de equidad (pues se dotaría de capital humano a las zonas que más lo necesitan), así como el de eficiencia

¹ Ver Button (1998).

(pues dichas inversiones obtendrían un mayor retorno). Así, la política educativa puede jugar un papel preponderante en favorecer la convergencia regional, utilizando al mismo tiempo de la forma más eficiente los recursos públicos disponibles. Sin embargo, se debe señalar que lo anterior sólo es válido si se acompaña de una política integral de desarrollo que fomente la inversión y el crecimiento, que genere incentivos para que el individuo permanezca en tales zonas; si no se crean los estímulos necesarios para que la productividad local del individuo sea comparable a aquella en zonas desarrolladas, sería más eficiente que él migre a zonas desarrolladas.

Dicho lo anterior, el objetivo de la presente investigación es estudiar cómo varían los rendimientos a la educación a través de las regiones geográficas en México. Para ello se estimarán ecuaciones *Mincerianas* por región, con distintas especificaciones econométricas; el ejercicio se realiza considerando los supuestos del modelo lineal clásico, pero también tratando el problema de endogeneidad de la educación.

El trabajo se divide de la siguiente manera. En la primera sección se reseñan brevemente algunos de los trabajos que han estudiado los rendimientos a la escolaridad en México. En las secciones 2 y 3 se presenta el modelo que se estimará y la descripción estadística de los datos, respectivamente. En el apartado 4 se ofrecen las estimaciones del modelo lineal clásico, mientras que en la parte 5 se presentan las estimaciones cuando se controla por el carácter endógeno de la educación.

1. REVISIÓN BIBLIOGRÁFICA

La teoría de capital humano afirma que la educación es una inversión que permite adquirir conocimientos y mayor productividad; de esta forma, la predicción empírica de dicha teoría es que un mayor nivel de escolaridad se verá acompañado de un mayor ingreso.

La teoría modela la elección del nivel de escolaridad como el resultado de la maximización del valor presente descontado de las ganancias de un individuo en el curso de su vida, hasta el momento en que se retira del mercado laboral. Los supuestos del modelo son que el individuo trabaja o estudia, pero no puede hacer ambas actividades al mismo tiempo, y que los costos directos de la educación son cero; el modelo también introduce una ecuación diferencial que representa cómo cambia el capital humano en el tiempo.² Las conclusiones del modelo son que el nivel de escolaridad aumenta con la duración de la vida activa del individuo y con la habilidad del mismo.

Con base en el modelo anterior, Mincer (1974) dedujo una forma funcional que relaciona los ingresos con la inversión en educación, y que permite calcular la tasa interna de retorno para ésta última variable. El autor demostró que dicha tasa puede ser interpretada como el incremento relativo del ingreso por un año adicional de escolaridad, cuando la ecuación es estimada por mínimos cuadrados ordinarios (MCO). El autor también supuso la posibilidad de adquirir conocimientos en el trabajo, por lo que la experiencia también tendría un efecto importante en los ingresos del individuo. De esta forma, en la ecuación propuesta por Mincer (1974), el ingreso individual (en logaritmos) se estima en función de la experiencia, la experiencia al cuadrado y los años de educación del individuo (en forma

² El modelo se presenta con detalle en Cahuc y Zylberberg (2004).

aditiva).³

El trabajo de Mincer (1974) ha generado una extensa investigación empírica a nivel mundial. Asimismo, existen trabajos que tratan de evaluar la vigencia de la ecuación *Minceriana*, o que ofrecen y estudian algunas sugerencias que permiten mejorar los alcances de la misma. Dos artículos que resumen estas extensiones metodológicas son Card (1999) y Lemieux (2006).

Card (1999) reconoce la importancia de la ecuación salarial *Minceriana* como un punto de partida para estudiar los efectos de la educación en el mercado laboral. El autor presenta un modelo simple estático que describe la decisión del individuo de educarse, y que permite explicar matemáticamente de dónde proviene el sesgo de habilidad del retorno a la educación al usar MCO.⁴ Posteriormente, estudia el uso de variables instrumentales (VI) para solucionar el problema de endogeneidad, el cual surge cuando la covarianza entre la variable educación y el término de error es diferente de cero. El autor profundiza en el uso de variables que representan los antecedentes familiares y los modelos para hermanos y gemelos, y menciona las condiciones bajo las cuales dichos esquemas generan estimadores consistentes de los retornos a la educación.

En un estudio más reciente, Lemieux (2006) se pregunta si la forma funcional básica de la ecuación *Minceriana* sigue siendo válida en la actualidad, o si es necesario revisar el modelo de referencia. En general el autor deduce que la ecuación *Minceriana*

³ Mincer (1974) probó también el efecto de la inclusión de un término cuadrático de la variable escolaridad y de la interacción de la escolaridad con la experiencia, bajo la idea de que el impacto de la educación puede variar con la duración de la misma. Sin embargo, para los datos con los que trabajó, dicho resultado no fue significativo cuando el ingreso se mide en ganancias semanales.

⁴ El sesgo de habilidad se refiere a la influencia de las capacidades individuales no observadas sobre el nivel de escolaridad alcanzado. Esto implica que en una ecuación salarial el supuesto de covarianza cero entre el error y la variable de escolaridad no se cumple, lo que hace que los retornos a la educación de MCO sean sesgados. Ver Cahuc y Zilberberg (2004).

permanece como una referencia adecuada para estimar ecuaciones salariales para el caso estadounidense, no obstante es preciso realizar algunos ajustes.⁵ Tales ajustes se refieren a la inclusión de la variable experiencia en un grado polinomial mayor (como experiencia a la potencia cuarta) con el objetivo de representar mejor el comportamiento de los salarios en el ciclo vital, e incluso considerar incluir la variable de escolaridad en forma cuadrática (si se confirma que la relación entre escolaridad y salarios es no lineal). Asimismo, el autor considera útil permitir efectos generacionales, bajo la justificación de que los retornos a la educación pueden diferir por grupos de edad debido a cambios en la estructura salarial.

1.1 Estudios para México

Para el caso de México existen varios trabajos que estudian la relación entre ingresos y educación; casi todos los estudios se han realizado a un nivel nacional, y han tomado como referencia alguna de las encuestas realizadas por el Instituto Nacional de Estadística, Geografía e Informática (INEGI). Las investigaciones se diferencian por los métodos de estimación utilizados, las muestras y la forma en que se corrigen diversos sesgos de estimación.

A continuación se reseñan algunas investigaciones realizadas para México. Estas se agruparon de la siguiente manera: las que estudian los retornos a la educación sin corregir por ningún sesgo; las que corrigen por algún sesgo de estimación; otra que prueba la existencia del credencialismo en México; y por último, un estudio de los retornos a la educación a nivel regional.

Bracho y Zamudio (1994) estiman una ecuación salarial *Minceriana*, para lo cual

⁵ Algunos de estos ajustes son mencionados brevemente por Card (1999).

utilizan los microdatos de la Encuesta Nacional de Ingreso y Gasto de los Hogares (ENIGH) 1989. Los autores calculan los retornos de la educación con la variable educación como variable continua y también con variables dicotómicas de nivel de escolaridad. Además las estimaciones se realizan para la muestra total, y distinguiendo por género, por zona de residencia (es decir, se diferencia entre zonas rurales y urbanas) y por efectos generacionales.⁶ Los autores encuentran un rendimiento promedio de la educación de 11.7 por ciento (es decir, un año adicional de educación genera un incremento porcentual de esa magnitud en el ingreso); sin embargo dicho retorno es mayor para los hombres, en las zonas rurales y para el grupo de edad joven.

Al estimar la ecuación salarial incluyendo las variables dicotómicas por nivel de educación, se observa un claro perfil creciente de los retornos. Cuando se divide la muestra por género, los autores encuentran que el perfil es creciente sólo para los hombres, mientras que la tasa de retorno de las mujeres muestra una disminución en el nivel superior. Asimismo, la zona urbana muestra un perfil creciente más pronunciado de los retornos, mientras que en la zona rural los retornos son más o menos constantes a excepción de la educación del nivel medio superior, con un retorno de 16.3 por ciento con respecto al nivel inmediato anterior (secundaria).

López (2001) realiza un estudio completo acerca de la situación de la distribución de ingresos en México en el periodo 1984-1996, además de dar una breve reseña de la evolución de los logros académicos del país. En su trabajo estima los retornos de la educación con una ecuación salarial *Minceriana* que incluye controles para el género del

⁶ Respecto a este último punto los autores dividen la muestra por dos grupos de edad: de 20 a 34 y de 35 a 49 años.

individuo, el estatus laboral, la actividad económica en que trabaja y la región geográfica.⁷ La encuesta utilizada es la Encuesta Nacional de Empleo Urbano (ENEU) para los años 1988, 1992, 1996 y 1997, y se efectúan estimaciones MCO y regresiones de cuantil.⁸

Debido a que la autora introduce los diferentes niveles educativos alcanzados como variables dicotómicas, estima el valor marginal de la educación por cada nivel de educación y cuantil.⁹ En general, se observa que los estimadores MCO son bastante similares a aquellos obtenidos para los cuantiles 0.5 y 0.75. Los resultados muestran que la educación juega un papel crucial en el proceso de formación de ingresos, que el efecto de dicha variable no es homogéneo a través de la distribución *condicional* de ingresos;¹⁰ y que el valor marginal de la educación no ha cambiado significativamente (en el tiempo) en la educación básica.

Con el objetivo de corregir por el sesgo de elección en la educación superior,¹¹ Zamudio (1995) estima los rendimientos educativos de este nivel en México, para lo cual utiliza el método de corrección por selectividad de Heckman. El autor comenta que en la educación superior el problema de endogeneidad reviste especial importancia, pues el acceso a este nivel educativo influyen características de los individuos y factores socioeconómicos; debido a que muchos de estos factores forman parte de la decisión de continuar con estudios superiores y del desempeño laboral del individuo, y no son observables, existe el riesgo de obtener estimadores sesgados en los retornos a la educación

⁷ Las regiones geográficas utilizadas son Norte, Centro, Sur y Ciudad de México.

⁸ Los cuantiles utilizados son 0.10, 0.25, 0.50, 0.75 y 0.90.

⁹ Las variables dicotómicas de educación son para primaria incompleta, primaria completa, secundaria completa, preparatoria completa y universidad completa.

¹⁰ La autora se refiere a la distribución de ingresos condicional a las variables de control (edad, edad al cuadrado, género, estatus laboral, sector económico y región).

¹¹ Se refiere al sesgo que resulta por no considerar a la escolaridad como una variable de elección para el individuo.

si se estiman ecuaciones vía MCO. Las ecuaciones de ingresos y elección de educación se hacen para los individuos con estudios máximos de preparatoria y para aquellos con educación superior; el autor utiliza la información de la ENIGH 1989.

Una vez obtenidas las ecuaciones de ingreso y escolaridad, el autor estima numéricamente las tasas de retorno, y observa que los resultados casi se triplican cuando se ajusta por sesgo. Cuando los retornos se calculan con los estimadores MCO sin ajustar por el sesgo, el retorno es de 10.5 por ciento, a diferencia de 28.6 y 26.9 por ciento si se ajusta por dicho sesgo. El autor advierte que los resultados deben tomarse con cuidado porque la muestra utilizada tiene muchas restricciones, y por lo tanto no pueden ser generalizados a toda la población.¹²

En la línea del trabajo anterior, y considerando el carácter endógeno de la educación, Barceinas (2003) presenta una serie de procedimientos para estimar los rendimientos de la educación en México; tales procedimientos se basan principalmente en la aplicación de métodos de VI. El autor utiliza los datos de las ENIGH 1994 y 1998. Las estimaciones se realizan con diferentes determinantes de la escolaridad: la edad, el producto interno bruto (PIB) real per cápita y el gasto real en educación per cápita en distintos momentos del ciclo vital del individuo,¹³ los antecedentes familiares (como el nivel de escolaridad o la ocupación de los padres) y con instrumentos de rango. En general el autor obtiene rendimientos más altos al estimar con VI que con MCO; la diferencia porcentual varía significativamente según el año y la forma de estimación de VI.

¹² Por ejemplo, el autor incluye en la muestra sólo a los individuos que estuvieran viviendo con sus padres al momento del levantamiento de la encuesta.

¹³ El autor utiliza el PIB real per-cápita y el gasto real en educación per cápita en distintos momentos del ciclo vital del individuo: en el año de nacimiento, cuando tiene 6 años y cuando tiene 12; el autor elige los mejores instrumentos para cada encuesta.

Bajo una perspectiva diferente respecto a los trabajos anteriormente reseñados, Guizar (2007) se pregunta si el salario de los individuos refleja verdaderamente su productividad marginal o si las personas se enfrentan a un esquema de señalización que asigna puestos de trabajo de manera discrecional. En este sentido, su trabajo consiste en estimar el impacto de las credenciales educativas sobre el salario en el mercado de trabajo en México, y corroborar la existencia de un efecto *sheepskin*.¹⁴

El autor estima una ecuación *Minceriana*, donde se incluyen además como variables explicativas dicotómicas que representan: los diplomas obtenidos por los individuos al finalizar los diferentes niveles educativos (primaria, secundaria, preparatoria, normal y posgrado), las credenciales técnicas para los individuos que terminaron los niveles anteriormente mencionados, el género del individuo, el estado civil, el número de hijos y si pertenecen a alguna etnia. El autor utiliza el método de Heckman para controlar el sesgo de autoselección, y la información la obtiene de la Encuesta Nacional de Niveles de Vida de los Hogares (ENNVIIH) 2002; dicha encuesta da cuenta de las credenciales y de los años de educación de forma separada.¹⁵

Los resultados muestran efectos positivos y significativos de los diplomas de secundaria, preparatoria y escuela normal -y de los títulos de licenciatura y posgrado- sobre el ingreso de las mujeres; y efectos positivos de los títulos de licenciatura y posgrado para los hombres. Los efectos *sheepskin* así encontrados parecen validar la hipótesis de credencialismo en el caso de México.

Por último, es importante señalar que ya existe un trabajo realizado para evaluar los

¹⁴ Efecto *sheepskin* se refiere a que los retornos privados a la educación se den con respecto a los certificados y no a los años de educación per-se.

¹⁵ La encuesta fue levantada conjuntamente por el CIDE y la UIA.

retornos a la educación a nivel regional. Barceinas (2005) estudia los rendimientos de la educación por región en México, e indaga en qué medida la distribución regional del nivel educativo en el país puede explicar la convergencia regional de los ingresos. El autor estima ecuaciones *Mincerianas* estándar regionales,¹⁶ para los años 1984-2002 para tratar de probar si los rendimientos a la educación tienden a ser más elevados en las regiones menos desarrolladas, tal y como ocurriría si el producto regional dependiese de una función de producción cóncava de la educación. La información proviene de las ENIGH, y considera a individuos con edad entre 16 y 65 años, que trabajaran tiempo completo (más de 35 horas a la semana) y que recibieran ingreso únicamente por remuneraciones al trabajo.

Barceinas (2005) encuentra que las regiones más pobres (sur y península) son las que obtienen los rendimientos más altos, mientras que las regiones más desarrolladas, que corresponden a las del norte del país, muestran rendimientos más bajos, en especial a partir de 1994. Al graficar los rendimientos de la educación contra el ingreso per-cápita o el nivel educativo, es posible apreciar una relación negativa.¹⁷ Se observa además un incremento en los rendimientos a la educación en el periodo 1984-1994; en este último año algunas regiones alcanzan su máximo, para después descender ligeramente o mantenerse en los mismos niveles hasta 2002.

Por último, es importante mencionar que el autor también estima en qué medida la

¹⁶ El autor utiliza diez regiones: Noroeste, Norte-Centro, Noreste, Centro-Norte, Centro-Oeste, Centro, Capital, Este, Sur y Península.

¹⁷ Para constatar de una manera más formal la relación negativa del rendimiento de la educación y el ingreso, por un lado, y la del rendimiento educativo con la escolaridad, por el otro, el autor realiza una estimación adicional, donde los rendimientos de la educación se estiman en función de variables dicotómicas por año y del ingreso anual o la escolaridad. De acuerdo al autor, los resultados confirman la asociación negativa entre las variables mencionadas; no obstante tal conclusión debe tomarse con cautela, debido a que en general los resultados de dicho ejercicio no son estadísticamente significativos.

desigual distribución de la escolaridad contribuye a explicar el proceso de divergencia (σ) regional. Para ello estima una ecuación *Minceriana* cuyo punto de referencia es el hogar y no el individuo, y realiza un ejercicio de simulación. Al comparar la desviación estándar observada del logaritmo natural de los ingresos familiares promedios entre regiones, con aquella obtenida del caso hipotético en el que la escolaridad promedio entre regiones es la misma, el autor encuentra que la desigualdad de los ingresos que se explica por la distribución heterogénea del capital humano varía de 25 hasta 43 por ciento, y dicho porcentaje parece ir a la par del ciclo económico del país.

El hecho de que la desigualdad del ingreso empeore cuando la economía marcha bien, y que la aportación que se explica por la desigual distribución del capital humano entre regiones se incremente, puede estar señalando que son las regiones más preparadas las que más beneficio obtienen de una expansión económica. Otra explicación alternativa es que en tiempos de crecimiento económico podría generarse un proceso de migración de los individuos más calificados hacia determinadas regiones, lo que justificaría la participación creciente de la educación en la explicación de la desigualdad regional.

Siguiendo el ejemplo de Barceinas (2005), el presente trabajo tiene como objetivo estimar los retornos a la educación en México a nivel regional, para 2005. A diferencia de ese trabajo previo, se introducirán cambios importantes que pretenden obtener estimadores más precisos; estos cambios se refieren al uso de una encuesta con representatividad estatal (y no nacional), al control de su diseño muestral, a la introducción de especificaciones econométricas más flexibles, y a la comparación estadística de los retornos a la educación. Así mismo se efectúan estimaciones por VI que permitirán apreciar en qué medida el carácter endógeno de la educación puede sesgar los estimadores de MCO.

2. EL MODELO

En el apartado anterior se describió brevemente la teoría del capital humano y cómo se derivó la ecuación *Minceriana*. En esta sección se presenta el modelo que se usará para obtener los retornos de la educación regionales para México; para ello se retoman algunas de las propuestas que pretenden mejorar los alcances de la ecuación tradicional.

El modelo se basa en la estimación de una ecuación *Minceriana* con algunos controles adicionales. En la versión más simple de la ecuación salarial propuesta por Mincer (1974), el logaritmo natural del salario del individuo es una función lineal en los parámetros de los años de escolaridad, la experiencia y la experiencia al cuadrado. La experiencia tiene un estatus ambiguo en la determinación del salario, pues ésta puede ser resultado no sólo de una inversión que absorbe parte del tiempo de trabajo (*aprender o hacer*) sino también del aprendizaje que la persona adquiere mientras trabaja (*aprender haciendo*). La forma en que la variable se introduce en la ecuación se basa en el supuesto de que el individuo adquiere una gran cantidad de conocimientos *adicionales* en el trabajo al inicio de su carrera, cuya tasa de adquisición disminuye con el tiempo.

La ecuación propuesta por Mincer (1974) se ha actualizado en su especificación para capturar cambios en la estructura de retornos salariales en años más recientes. Como se mencionó en la primera sección, Lemieux (2006) propone una serie de modificaciones con las cuales la ecuación *Minceriana* estándar puede explicar mejor empíricamente la relación ingreso-educación para el caso estadounidense. Tales cambios se refieren a la introducción de las variables de escolaridad y experiencia en grados polinomiales mayores.

Lemieux (2006) menciona diferentes razones por las cuales incluir la variable de escolaridad de forma lineal puede ser inadecuado. Por ejemplo, el logaritmo de los ingresos

puede ser una función cóncava de los años de escolaridad, si los individuos tienen diferentes preferencias (tasas de descuento) pero todos enfrentan una misma función producción cóncava, donde los retornos a la educación disminuyen conforme los años de escolaridad aumentan.¹⁸

Basado en sus estimaciones para la economía estadounidense, Lemieux (2006) concluye que incluir la escolaridad en forma lineal sólo podrá ser adecuado en un ambiente estable donde el crecimiento de la demanda relativa por trabajo calificado está aparejado por un crecimiento de la oferta relativa.

Por otra parte, el autor también realiza estimaciones para estudiar si la inclusión de la variable de experiencia al cuadrado captura bien el perfil de ingresos-experiencia. El autor encuentra que la especificación cuadrática subestima el crecimiento de los salarios para los primeros 10 a 15 años de experiencia; además dicha especificación también predice un descenso en los salarios pasados 25 años de experiencia, cuando en realidad los ingresos son muy estables entre los 25 y 40 años de experiencia. Esto no ocurre si se utiliza una especificación con un polinomio cuártico de la experiencia, el cual parece capturar muy bien las características del perfil ingresos-experiencia para Estados Unidos.

Aunque se han realizado muchos estudios para estimar los retornos de la educación en México, no existe alguno como aquel de Lemieux (2006) que examine si la versión simple de la ecuación *Minceriana* ajusta bien el perfil de ingresos-educación o ingresos-experiencia de la población mexicana. En vista de lo anterior, el presente trabajo se limita a buscar la mejor especificación de la ecuación salarial para cada región, con base en las

¹⁸ Mincer (1974) ya había considerado la posibilidad de que el retorno a la educación no fuera constante, sino que variara con el grado de escolaridad. Véase *supra*, p. 4, n. 3. Adicionalmente Mincer (1997) demostró que si hay heterogeneidad en las preferencias y oportunidades de ingreso de los individuos, el logaritmo de los ingresos puede ser una función cóncava o convexa de los años de escolaridad

extensiones tratadas por Lemieux (2006). El modelo econométrico general se describe a continuación:

$$\ln(\text{salhr}_{ir}) = \beta_{0r} + \beta_{1,r} \text{esc}_{ir} + \beta_{2,r} \text{esc}_{ir}^2 + \gamma_{1,r} \text{exper}_{ir} + \dots + \gamma_{4,r} \text{exper}_{ir}^4 + \delta_{1,r} \text{pibepc}_{ir} + \varepsilon_{ir} \quad (1)$$

$$i = 1, \dots, Nr; \quad r = 1, \dots, R$$

Donde:

- *salhr* representa el salario por hora individual.
- *esc* son los años de escolaridad; se incluyen de forma lineal, o como un polinomio cuadrático.
- *exper* son los años de experiencia, los cuales pueden introducirse de forma cuadrática o como un polinomio de orden mayor. Los años de experiencia (potencial) se definen como la edad menos los años de educación menos seis.
- *pibepc* es el PIB per cápita estatal de 2004, a precios de 1980.
- *i* representa al individuo, *r* a la región y *Nr* es el número de individuos en cada región.

Aún cuando incluir un intercepto regional permite controlar por la heterogeneidad existente entre regiones, se decidió incluir la variable del PIB estatal per cápita para controlar además por las heterogeneidades presentes al interior de cada región (es decir, entre entidades federativas).

El uso de la experiencia potencial ha sido práctica común en la mayoría de los estudios dedicados a estudiar los retornos a la educación, debido a que generalmente no se cuenta con los años reales de experiencia individual. Sin embargo, es importante tener en mente que, debido a que la experiencia real del individuo puede diferir de dicha medida, y que por ende la experiencia potencial está sujeta a errores de medición, el estimador de

MCO de la variable es inconsistente; los errores de medición hacen que el estimador de MCO sea atenuado y pueden inflar su varianza. No obstante, y a falta de los años de experiencia reales, se decidió usar tal medida como una variable *proxy*.

Como se mencionó anteriormente, la educación y la experiencia se incluirán en los polinomios que mejor ajusten a cada región. Las especificaciones así empleadas permitirán apreciar las diferencias regionales existentes en la magnitud de los retornos educacionales, pero también en la manera en que estos se comportan conforme el grado de escolaridad aumenta.¹⁹ Como una primera aproximación, los rendimientos a la educación se estimarán por MCO; posteriormente, la investigación ahondará en la obtención de tales rendimientos cuando se considera el carácter endógeno de la educación.

¹⁹ Aunque es claro que esto no sustituye la realización de estudios más profundos que investiguen formalmente los perfiles de ingreso-educación-experiencia para México.

3. ESTADÍSTICAS DESCRIPTIVAS

En este apartado se describe brevemente la base de datos a utilizar, la forma en que algunas de las variables fueron construidas y la estadística básica de las variables.

La base de datos utilizada proviene de la Encuesta Nacional de Ocupación y Empleo (ENOE), del Instituto Nacional de Estadística, Geografía e Informática (INEGI). El esquema de la encuesta es probabilístico, bietápico, estratificado y por conglomerados; tiene como unidad última de selección las viviendas particulares y como unidad de observación a las personas.

La razón principal por la cual se decidió usar la ENOE es que ésta es representativa a nivel estatal,²⁰ lo cual es importante dado el carácter regional de la investigación. La encuesta proporciona información básica del trabajador, como su ingreso, la actividad económica en que labora, lugar de residencia, y sus características personales como escolaridad, edad, estado civil y género, entre otras.

La ENOE reporta el ingreso laboral mensual que percibe el individuo y las horas semanales que trabaja, por lo que es posible calcular el salario por hora. Respecto a la escolaridad, la encuesta pregunta sobre el grado de escolaridad alcanzado (primaria, secundaria, preparatoria, etc.) y los años aprobados dentro de cada nivel; a partir de esta información se aproximaron los años de escolaridad totales por individuo.

Debido a que no se cuenta con una medición directa de los años de experiencia, ésta fue calculada como la edad menos los años de escolaridad menos seis. Dicha variable trata de capturar el perfil de ciclo de vida de ingresos, y el diferencial en la pendiente de esos perfiles a través de grupos de personas con diferentes grados de escolaridad.

²⁰ A diferencia de la ENIGH, que sólo tiene representatividad a nivel nacional.

El cuadro 1 presenta las estadísticas descriptivas básicas de las variables a utilizar, por región.²¹ En el Anexo 1 se muestra cómo se conforma cada zona.²² Las muestras se componen únicamente por hombres, con edad entre 20 y 60 años, y que recibieron ingresos principalmente por sueldo, salario o jornal. Las razones principales por las cuales se efectuaron dichos filtros son: como medio para reducir el problema de autoselección de entrar al mercado laboral (condicional a que el salario esperado exceda el salario de reserva), y que presumiblemente es más acentuado en muestras que incluyen mujeres; para incorporar a la población que más probablemente se encuentra en el mercado laboral; y para capturar de una manera más pura el efecto de la educación (y la productividad) en el ingreso, debido a que otro tipo de percepciones (como las financieras) pueden no estar ligadas directamente a la formación (y desempeño) del individuo. La información corresponde al tercer trimestre de 2005, por ser el trimestre que presenta menores variaciones estacionales año con año.

²¹ Las estadísticas y las estimaciones consideran el diseño muestral de la ENOE; de esta manera se controla por unidad primaria de muestreo (UPM), estrato y factor de expansión.

²² La regionalización se basa en Hanson & Woodruff (2003).

Cuadro 1: Estadísticas básicas de las variables, por región

Región/estadístico	Salario por hora	Edad	Años de escolaridad	Experiencia
<i>Frontera</i>				
Media	28.237	35.098	9.437	19.661
Desviación estándar	0.476	0.177	0.085	0.199
N: 7,037				
<i>Norte</i>				
Media	23.777	35.433	9.032	20.401
Desviación estándar	0.508	0.190	0.115	0.244
N: 6,645				
<i>Centro</i>				
Media	22.186	35.459	8.194	21.265
Desviación estándar	0.429	0.164	0.106	0.202
N: 11,860				
<i>Capital</i>				
Media	25.708	35.530	9.644	19.886
Desviación estándar	0.725	0.253	0.124	0.269
N: 2,959				
<i>Sureste</i>				
Media	22.542	35.341	8.884	20.457
Desviación estándar	0.646	0.190	0.133	0.241
N: 4,964				
<i>Sur</i>				
Media	19.185	35.034	8.508	20.526
Desviación estándar	0.531	0.283	0.202	0.340
N: 3,598				

N es el número de observaciones.

Fuente: Cálculos propios con información de la ENOE, tercer trimestre de 2005

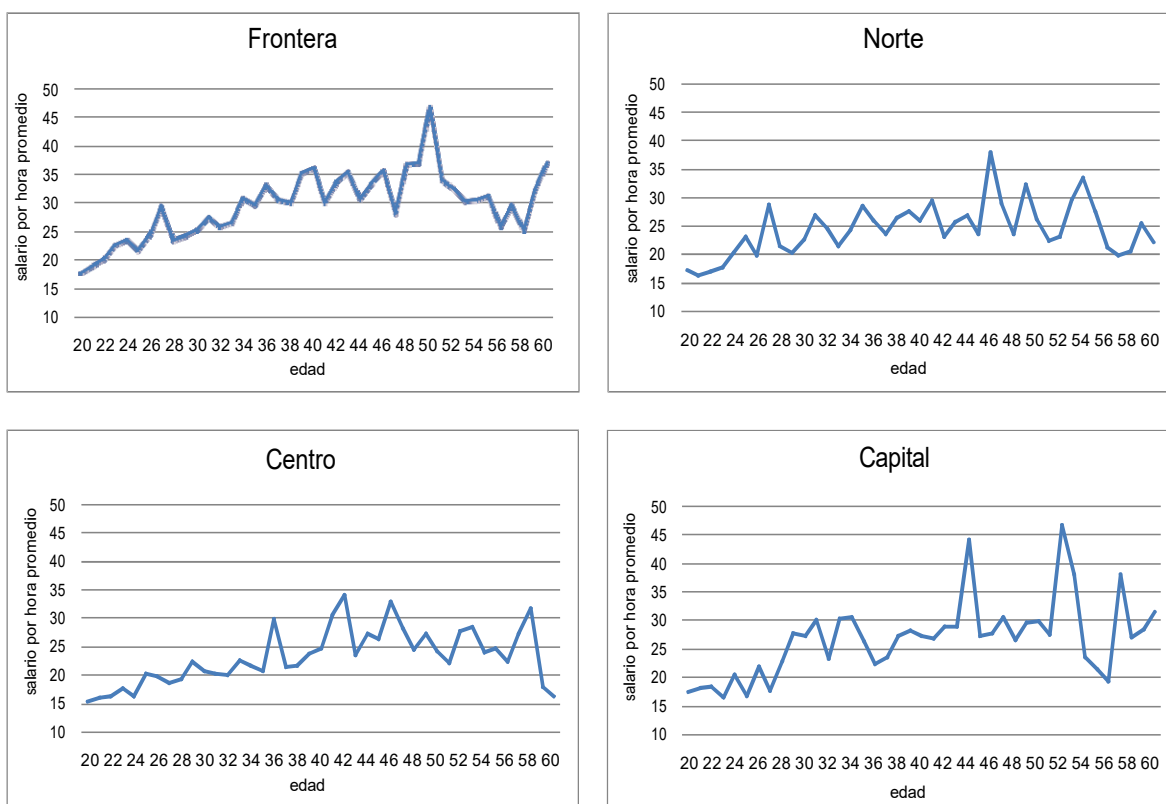
Lo primero a observar en el cuadro 1 son las pequeñas diferencias regionales en la edad promedio, y en menor medida, en la experiencia; esto no ocurre en los salarios y en la escolaridad, que presentan diferencias interregionales más notables.²³ Las regiones de la frontera, norte y capital se caracterizan por tener los salarios por hora y educación promedio más altos, en contraposición con las regiones del sur, sureste y centro. Estos dos indicadores son una pequeña muestra de las diferencias regionales en los niveles de

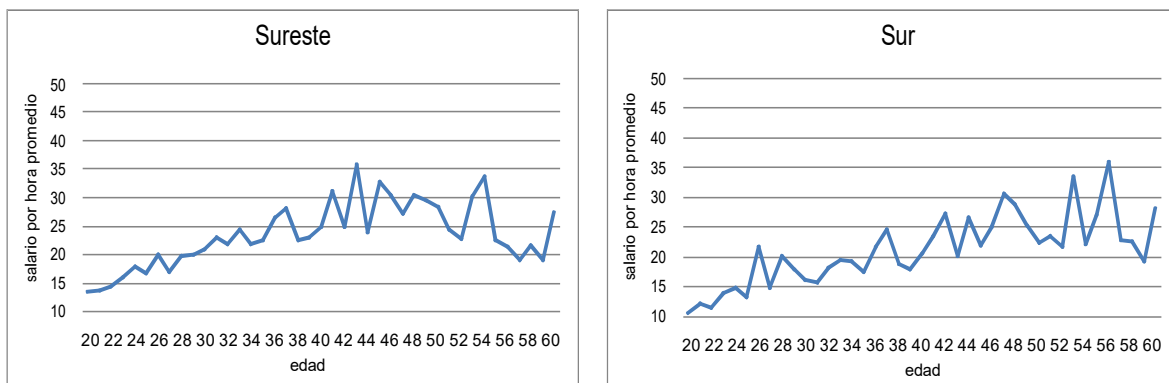
²³ Se calcularon intervalos de confianza para comprobar si las medias de las distintas variables difieren entre regiones; como tales intervalos no se traslapan, es posible afirmar que los promedios difieren estadísticamente entre las zonas. Ver Anexo 2.

desarrollo económico.

Las gráficas 1 muestran la relación entre el salario por hora promedio y la edad para cada región. En general, en todas las regiones se advierte un patrón creciente en los salarios por hora hasta alcanzar un máximo entre los 42 y 56 años de edad; después de esta edad los salarios promedio por hora en general disminuyen. Si se ajustara una línea de tendencia, en todos los casos se obtendría una relación cóncava, lo que es congruente con la predicción de la teoría respecto al comportamiento de los ingresos en el ciclo vital.

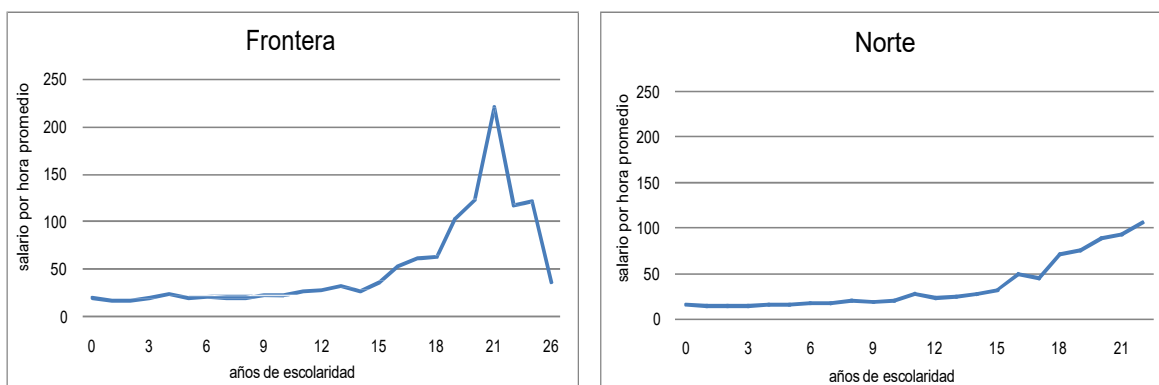
Gráficas 1: Relación ingresos-edad

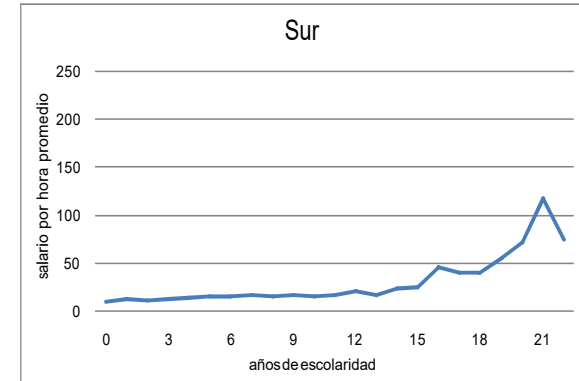
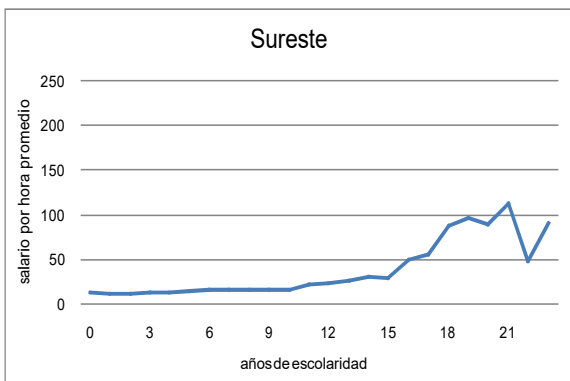
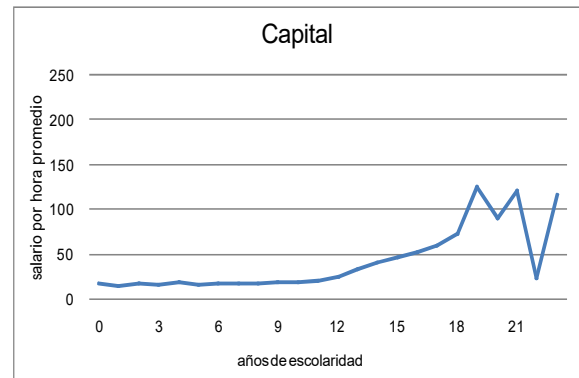
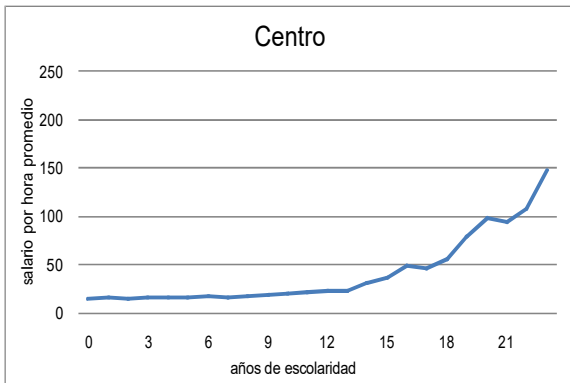




Las gráficas 2 muestran la relación entre el salario por hora promedio y los años de escolaridad para cada región. Las regiones muestran un comportamiento muy similar: los ingresos mantienen una tendencia ligeramente creciente en los primeros años de educación; al pasar los 12 años de escolaridad el crecimiento del salario aumenta, y el ingreso promedio por hora más alto se alcanza a los 20-22 años de formación. La relación ingresos-educación indica una relación convexa entre esas dos variables, al menos hasta los 21 años de escolaridad.

Gráficas 2: Relación ingresos- educación





4. RESULTADOS DEL MODELO LINEAL CLÁSICO

En este apartado se presentan los resultados de la estimación por MCO del modelo econométrico propuesto en la sección 2, y algunas relaciones que permitirán hacer algunas inferencias generales acerca de los retornos a la educación en el ámbito regional en México.

Con el objetivo de comparar estadísticamente si los retornos a la educación difieren de una región a otra, la estimación se realizó utilizando sólo una ecuación. Para ello se crearon variables dicotómicas que identifican cada una de las seis regiones, las cuales se interactuaron con las variables de años de escolaridad y experiencia; de esta manera se obtuvieron de una sola vez los parámetros correspondientes a cada zona, ver ecuación (1').

$$\ln(\text{salhr}_i) = d_r + \beta_1 \text{esc}_i \bullet d_r + \beta_2 \text{esc}_i^2 \bullet d_r + \gamma_1 \text{exper}_i \bullet d_r + \dots + \gamma_4 \text{exper}_i^4 \bullet d_r + \delta_1 \text{pibepc}_i \bullet d_r + \varepsilon_{i,r} \quad (1')$$

$$d_r = 1 \text{ ó } 0; r = 1, \dots, R$$

Se presentan cuatro especificaciones diferentes: la primera corresponde a la ecuación *Minceriana* tradicional; la segunda es la ecuación básica que incorpora la variable experiencia en polinomios mayores; la tercera añade a la especificación anterior la escolaridad al cuadrado; en la última, la introducción de polinomios mayores para la variable de educación o experiencia se realizó incluyendo sólo aquellos términos que fueron estadísticamente significativos.

Cabe mencionar que al controlar por el diseño muestral, se asume que los residuos de las estimaciones son independientes entre estratos, pero están idénticamente distribuidos en las UPM's con varianza común. El software utilizado (*Stata*) permite obtener errores

estándar robustos; la fórmula de la varianza del estimador MCO que se usa no impone restricciones sobre la forma de la heterocedasticidad y la correlación dentro de una UPM.²⁴

Los resultados de la estimación se presentan en el Cuadro 2.

²⁴ Ver Cameron & Trivedi (2005), capítulo 24.

Cuadro 2: Estimación del modelo lineal
variable dependiente: logaritmo natural del salario por hora
(errores estándar entre paréntesis)

Región/Variable	1	11	111	1V
Frontera				
Constante	1.7248 *** (0.0681)	1.7226 *** (0.0871)	2.1219 *** (0.0884)	2.1825 *** (0.0734)
Escolaridad	0.0878 *** (0.0028)	0.0882 *** (0.0029)	-0.0295 *** (0.0079)	-0.0303 *** (0.0076)
Escolaridad ²			0.0058 *** (0.0004)	0.0058 *** (0.0004)
Experiencia	0.0288 *** (0.0027)	0.0207 (0.0129)	0.0464 *** (0.0128)	0.0337 *** (0.0025)
Experiencia ²	-0.0003 *** (0.0001)	0.0008 (0.0010)	-0.0013 (0.0010)	-0.0005 *** (5.28E-05)
Experiencia ³		-4.71E-05 * (2.86E-05)	1.72E-05 (2.87E-05)	
Experiencia ⁴		5.57E-07 ** (2.74E-07)	-1.13E-07 (2.76E-07)	
P1BE per cápita	0.0016 *** (0.0005)	0.0016 *** (0.0005)	0.0020 *** (0.0005)	0.0019 *** (0.0005)
Norte				
Constante	1.5635 *** (0.0642)	1.4276 *** (0.0775)	1.7862 *** (0.0776)	1.7862 *** (0.0776)
Escolaridad	0.0820 *** (0.0029)	0.0834 *** (0.0030)	-0.0228 *** (0.0079)	-0.0228 *** (0.0079)
Escolaridad ²			0.0053 *** (0.0004)	0.0053 *** (0.0004)
Experiencia	0.0229 *** (0.0025)	0.0501 *** (0.0131)	0.0705 *** (0.0130)	0.0705 *** (0.0130)
Experiencia ²	-0.0002 *** (0.0001)	-0.0018 * (0.0010)	-0.0035 *** (0.0010)	-0.0035 *** (0.0010)
Experiencia ³		2.75E-05 (2.84E-05)	7.78E-05 *** (2.84E-05)	7.78E-05 *** (2.84E-05)
Experiencia ⁴		-1.26E-07 (2.72E-07)	-6.52E-07 ** (2.74E-07)	-6.52E-07 ** (2.74E-07)
P1BE per cápita	0.0046 *** (0.0006)	0.0046 *** (0.0006)	0.0050 *** (0.0006)	0.0050 *** (0.0006)

Cuadro 2: Estimación del modelo lineal (continuación)

Región/Variable	1	11	111	1V
Centro				
Constante	1.3019 *** (0.0721)	1.2201 *** (0.0887)	1.4404 *** (0.0936)	1.3019 *** (0.0721)
Escolaridad	0.0805 *** (0.0028)	0.0811 *** (0.0028)	-0.0059 (0.0083)	0.0805 *** (0.0028)
Escolaridad ²			0.0047 *** (0.0004)	
Experiencia	0.0300 *** (0.0026)	0.0475 *** (0.0133)	0.0742 *** (0.0130)	0.0300 *** (0.0026)
Experiencia ²	-0.0004 *** (0.0001)	-0.0015 (0.0010)	-0.0035 *** (0.0010)	-0.0004 *** (0.0001)
Experiencia ³		2.52E-05 (2.98E-05)	0.0001 *** (2.91E-05)	
Experiencia ⁴		-1.87E-07 (2.91E-07)	-7.60E-07 *** (2.84E-07)	
P1BE per cápita	0.0094 *** (0.0009)	0.0094 *** (0.0009)	0.0097 *** (0.0009)	0.0094 *** (0.0009)
Capital				
Constante	1.6359 *** (0.0682)	1.5773 *** (0.1090)	2.2079 *** (0.1087)	2.2079 *** (0.1087)
Escolaridad	0.0968 *** (0.0044)	0.0975 *** (0.0044)	-0.0601 *** (0.0107)	-0.0601 *** (0.0107)
Escolaridad ²			0.0076 *** (0.0005)	0.0076 *** (0.0005)
Experiencia	0.0222 *** (0.0040)	0.0312 (0.0209)	0.0582 *** (0.0205)	0.0582 *** (0.0205)
Experiencia ²	-0.0002 ** (0.0001)	-0.0005 (0.0015)	-0.0029 ** (0.0015)	-0.0029 ** (0.0015)
Experiencia ³		-5.55E-07 (4.26E-05)	7.68E-05 * (4.11E-05)	7.68E-05 * (4.11E-05)
Experiencia ⁴		8.76E-08 (4.02E-07)	-7.70E-07 ** (3.86E-07)	-7.70E-07 ** (3.86E-07)
P1BE per cápita	0.0007 *** (0.0002)	0.0007 *** (0.0002)	0.0007 *** (0.0002)	0.0007 *** (0.0002)

Cuadro 2: Estimación del modelo lineal (continuación)

Región/Variable	1	11	111	1V
Sureste				
Constante	1.1807 *** (0.0589)	0.9800 *** (0.0854)	1.4208 *** (0.0861)	1.4208 *** (0.0861)
Escolaridad	0.1010 *** (0.0032)	0.1022 *** (0.0032)	-0.0177 ** (0.0088)	-0.0177 ** (0.0088)
Escolaridad ²			0.0062 *** (0.0004)	0.0062 *** (0.0004)
Experiencia	0.0367 *** (0.0033)	0.0783 *** (0.0161)	0.0969 *** (0.0157)	0.0969 *** (0.0157)
Experiencia ²	-0.0004 *** (0.0001)	-0.0029 ** (0.0012)	-0.0045 *** (0.0012)	-0.0045 *** (0.0012)
Experiencia ³		4.90E-05 (3.47E-05)	9.67E-05 *** (3.45E-05)	9.67E-05 *** (3.45E-05)
Experiencia ⁴		-2.92E-07 (3.35E-07)	-7.88E-07 ** (3.34E-07)	-7.88E-07 ** (3.34E-07)
P1BE per cápita	0.0031 *** (0.0005)	0.0031 *** (0.0005)	0.0030 *** (0.0004)	0.0030 *** (0.0004)
Sur				
Constante	0.7376 *** (0.1397)	0.4895 *** (0.1628)	0.6100 *** (0.1681)	0.6100 *** (0.1681)
Escolaridad	0.0986 *** (0.0033)	0.1002 *** (0.0034)	0.0405 *** (0.0119)	0.0405 *** (0.0119)
Escolaridad ²			0.0032 *** (0.0006)	0.0032 *** (0.0006)
Experiencia	0.0321 *** (0.0037)	0.0953 *** (0.0194)	0.1104 *** (0.0191)	0.1104 *** (0.0191)
Experiencia ²	-0.0003 *** (0.0001)	-0.0049 *** (0.0015)	-0.0062 *** (0.0015)	-0.0062 *** (0.0015)
Experiencia ³		0.0001 *** (4.27E-05)	0.0002 *** (4.32E-05)	0.0002 *** (4.32E-05)
Experiencia ⁴		-1.09E-06 *** (4.12E-07)	-1.46E-06 *** (4.22E-07)	-1.46E-06 *** (4.22E-07)
P1BE per cápita	0.0186 *** (0.0043)	0.0186 *** (0.0043)	0.0198 *** (0.0042)	0.0198 *** (0.0042)
R ²	0.9665	0.9666	0.9687	0.9682
No. Observaciones	37,063	37,063	37,063	37,063
Estadístico F	13,004.63	9,518.50	9,175.12	9,892.06
Estadístico F de H1	8.27 ***	8.41 ***	8.82 ***	96.65 ***
Estadístico F de H2			8.13 ***	8.80 ***

* Significativo al 90%; ** significativo al 95%; *** significativo al 99%; sin asterisco: no significativo

H1: los coeficientes de escolaridad son iguales entre las regiones

H2: los coeficientes de escolaridad al cuadrado son iguales entre regiones

Fuente: Cálculos propios con información de la ENOE, tercer trimestre de 2005.

En general los signos observados son los que se esperaban. Sin embargo, al considerar la especificación IV, se observa que, a excepción de la región centro, el retorno a la educación no es constante conforme el grado de escolaridad aumenta (pues el coeficiente de la escolaridad al cuadrado es estadísticamente significativo). Además, sin contar a las regiones centro y frontera, los polinomios mayores de la experiencia contribuyen a explicar los salarios de los individuos. Las pruebas F reportadas al pie del cuadro 2 indican que los coeficientes de escolaridad son estadísticamente diferentes entre regiones, a un nivel de confianza mayor al 99 por ciento.

Para el cálculo de los efectos marginales de la educación, se utiliza la educación promedio nacional.²⁵ Los retornos a la escolaridad así calculados señalan que las regiones con mayores rendimientos a la educación son la sur y la sureste, así como la centro, con retornos de 9.86, 9.32 y 8.05 por ciento, respectivamente. En contraste, las regiones con menores rendimientos a la escolaridad son la capital, la frontera y el norte, con 7.58, 7.40 y 7.25 por ciento. Los efectos marginales son estadísticamente diferentes entre sí, lo cual se comprueba al observar que los intervalos de confianza no se traslapan entre regiones; ver Anexo 3. En todos los casos (a excepción del centro), el signo de la escolaridad al cuadrado indica que la relación entre años de educación y el ingreso es convexa, es decir, mayores años de educación implican retornos cada vez más altos.²⁶

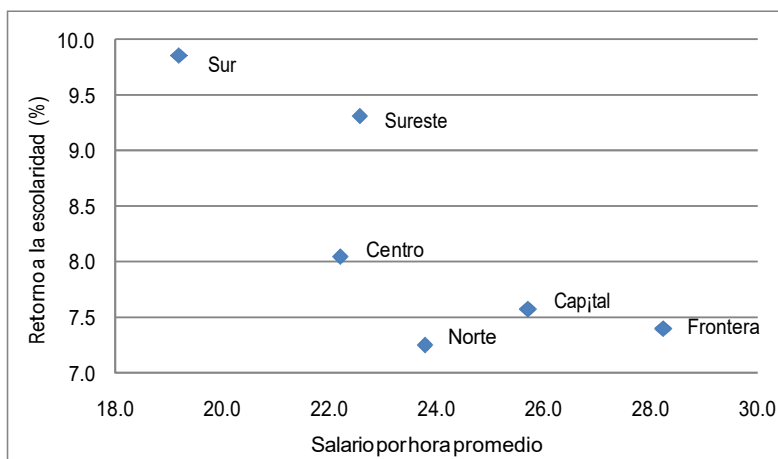
Si se relacionan los retornos obtenidos con la escolaridad y el ingreso promedio, es posible advertir cierta tendencia negativa; las regiones con menor nivel educacional y

²⁵ De esta manera, el retorno a la escolaridad para una región r es: $retorno_r = \beta_1 + 2*\beta_2*(escolaridad\ promedio\ nacional)$, donde β_1 y β_2 son los coeficientes de la escolaridad en forma lineal y cuadrática, respectivamente.

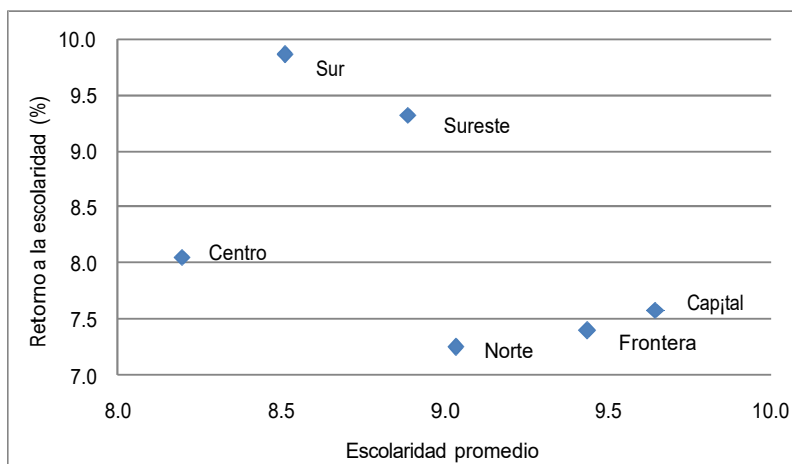
²⁶ Al tomar la especificación IV, los años de escolaridad a partir de los cuales un año más de estudios genera un retorno positivo son el 3ro, 2do, 4to y 1ro, respectivamente, para las regiones frontera, norte, capital y sureste. En la región sur el retorno siempre es creciente, mientras que en el centro el retorno es constante, como se aprecia en el cuadro 1.

salario promedio son las que parecen tener asociado un mayor retorno a la educación (ver gráficas 3 y 4). Estos resultados confirman aquellos encontrados por Barceinas (2005).

Gráfica 3. Relación de ingreso y retorno a la escolaridad



Gráfica 4. Relación de años de educación y retorno a la escolaridad



Las estimaciones presentadas en este trabajo son evidencia de retornos a la educación diferenciados al interior del país. Distintas especificaciones y pruebas estadísticas indican que tales rendimientos difieren en su magnitud, y que a excepción de

un solo caso, el retorno no es constante, sino creciente (al menos hasta los 21 años de escolaridad, como se aprecia en las gráficas 2).²⁷

Por último es importante recordar que si el término de error de la ecuación 1 está correlacionado con la variable educación, la estimación de MCO genera estimadores sesgados e inconsistentes. En este sentido, sería útil controlar por el carácter endógeno de la educación, y de esta manera comprobar si la relación entre nivel de ingresos-educación y retornos a la escolaridad se sigue manteniendo.

²⁷ Una especificación que incluyera variables dicotómicas por nivel o grado de escolaridad permitiría controlar por lo que parece una caída en el retorno a la educación después de los 21 años de escolaridad, como se observa en las gráficas 2.

5. RESULTADOS DEL MODELO DE VARIABLES INSTRUMENTALES

Con el objetivo de observar en qué magnitud los retornos a la educación cambian cuando se considera el carácter endógeno de la educación, en esta sección se presentan las estimaciones con VI de las ecuaciones salariales *Mincerianas*. El método utilizado es mínimos cuadrados en dos etapas (MC2E).

Los instrumentos para los años de escolaridad son el gasto real en educación per cápita en distintos momentos de la vida del individuo (0, 6 y 12 años de edad), y una variable dicotómica que distingue a los individuos a los cuales la reforma educativa de 1993 pudo haber afectado.

La utilización del gasto real per cápita en educación sigue la línea del trabajo previo de Barceinas (2003). La variable trata de capturar cómo la cantidad de recursos destinados a la educación (y por lo tanto, la disponibilidad de servicios educativos), puede influir en la decisión de escolaridad del individuo, y en qué momento de la vida del individuo dicha variable tiene un mayor impacto. El gasto en educación tiene una importancia especial porque en México la mayor parte de los servicios educativos son financiados por los gobiernos estatales y el federal.²⁸

La obligatoriedad de la educación también puede tener un impacto significativo en la escolaridad de las personas. La Ley General de Educación en México establece que el Estado está obligado a prestar servicios educativos para que toda la población pueda cursar la educación preescolar, primaria y secundaria, aunque solamente estos dos últimos niveles tienen un carácter obligatorio. La última reforma en materia de obligatoriedad, realizada en

²⁸ Información de la Secretaría de Educación Pública (SEP) señala que, para el ciclo 2005-2006, el 71.8 por ciento de los alumnos asiste a escuelas administradas por los gobiernos estatales, mientras que 10.4 por ciento asiste a escuelas administradas por la Federación.

1993, fue la que dio a la educación secundaria su condición actual. Es posible que las

personas que al momento de la reforma estuvieran terminando la primaria o estudiando algún grado menor, tengan mayores años de educación con respecto las personas que en 1993 ya habían terminado sus estudios. En este orden de ideas se creó una variable dicotómica que distingue aquellas personas nacidas en 1981 o después.

Al considerar que la experiencia es una variable potencialmente endógena (pues se calcula como la edad menos los años de educación menos 6), se decidió considerarla también como tal; para ello se asumió que la experiencia se explica por la edad del individuo (en un polinomio de orden dos o cuatro, si la especificación incluye un polinomio cuadrático o cuarto de la experiencia, respectivamente), y por el gasto en educación real per cápita en forma lineal y cuadrática (para las tres edades consideradas). Por último, el ingreso del individuo se estima en función de la escolaridad (en forma lineal), del PIB estatal per-cápita y de la experiencia en un polinomio de grado dos o cuatro. Dicho lo anterior, las ecuaciones a estimar son:

$$esc_{ir} = \delta_0 + \delta_{1,r} gerpc0_{ir} + \delta_{2,r} gerpc6_{ir} + \delta_{3,r} gerpc12_{ir} + \delta_{4,r} reforma93_{ir} + \vartheta_{ir}$$

$$\begin{aligned} (exper_{ir})^n = & \varphi_0 + \sum_{j=1}^n \zeta_j (edad_{ir})^j + \phi_1 gerpc0 + \phi_2 gerpc6 + \phi_3 gerpc12 \\ & + \phi_4 (gerpc0)^2 + \phi_5 (gerpc6)^2 + \phi_6 (gerpc12)^2 + \omega_{ir} \end{aligned} \quad (2)$$

$$\ln(salhr_{ir}) = \beta_0 + \beta_{1,r} esc_{ir} + \gamma_{1,r} exper_{ir} + \dots + \gamma_{4,r} exper_{ir}^4 + \delta_1 pibepc_{ir} + \varepsilon_{ir}$$

$$i = 1, \dots, Nr; \quad r = 1, \dots, R$$

Donde:

- $gerpc0$, $gerpc6$ y $gerpc12$ es el gasto en educación real per cápita cuando el

individuo tenía 0, 6 y 12 años de edad.²⁹

- *reforma93* es una variable dicotómica que toma el valor de 1 para las personas que nacieron en 1981 o después.
- *esc*, *salhr* y *pibepc* son la escolaridad, el salario por hora del individuo y el PIB estatal per-cápita, tal como se definieron con anterioridad. edad se refiere a la edad del individuo.

Las ecuaciones (2) cumplen así con la condición de orden de identificación, dado que el sistema está sobreidentificado.

Es claro que los antecedentes familiares juegan un papel determinante en la educación del individuo. En ese aspecto, sin embargo, el alcance de la ENOE, al igual que el de la mayoría de las encuestas realizadas en México, es limitado, pues no es posible obtener las características de los padres a menos que los individuos vivan aún con sus progenitores. Si se quisiera utilizar alguna característica de los padres (como su educación) para explicar la escolaridad individual, la muestra se reduciría sólo a personas que viven con sus padres, lo que genera dudas sobre su aleatoriedad; por ello se decidió sacrificar este tipo de variables. Sin embargo, se debe tener en mente que omitir esos regresores tampoco está exento de problemas; si las variables de gasto real en educación per cápita, de la reforma educativa y la edad explican sólo débilmente la escolaridad y la experiencia, VI puede generar estimadores sesgados e ineficientes; estas deficiencias no desaparecen aún cuando se tengan muestras muy grandes.³⁰

El cuadro 3 presenta las estimaciones de MCO y VI para dos especificaciones

²⁹ Las series están a precios de 1980, y fueron construidas a partir de información de INEGI, Banxico y Conapo.

³⁰ Ver Wooldridge (2002), Cap.5.

distintas.³¹ La forma de estimación difiere respecto a la utilizada en la sección 4, en el sentido de que por cada especificación, en esta parte se realizan seis estimaciones diferentes, una por región; implícitamente se asume que las regiones tienen varianzas distintas. El objetivo de esta sección es sólo aportar luz sobre el grado en que la endogeneidad de la educación podría estar sesgando las estimaciones de MCO.

³¹ Aunque se realizaron estimaciones VI incluyendo el término cuadrático de la educación, no se obtuvieron parámetros estadísticamente significativos.

Cuadro 3: Estimación del modelo lineal, mínimos cuadrados ordinarios
y variables instrumentales

variable dependiente: logaritmo natural del salario por hora
(errores estándar entre paréntesis)

Región/Variable	1		11	
	MCO	V1	MCO	V1
Frontera				
Constante	1.7248 *** (0.0681)	2.1052 *** (0.2724)	1.7226 *** (0.0871)	1.5365 *** (0.3789)
Escolaridad	0.0878 *** (0.0028)	0.0392 (0.0325)	0.0882 *** (0.0029)	0.0901 ** (0.0403)
Experiencia	0.0288 *** (0.0027)	0.0346 *** (0.0045)	0.0207 (0.0129)	0.0663 *** (0.0258)
Experiencia ²	-0.0003 *** (0.0001)	-0.0005 *** (0.0001)	0.0008 (0.0010)	-0.0027 (0.0021)
Experiencia ³			-4.71E-05 * (2.86E-05)	5.27E-05 (6.07E-05)
Experiencia ⁴			5.57E-07 ** (2.74E-07)	-3.87E-07 (5.94E-07)
P1BE per cápita	0.0016 *** (0.0005)	0.0022 *** (0.0007)	0.0016 *** (0.0005)	0.0016 ** (0.0007)
N= 6877				
Norte				
Constante	1.5635 *** (0.0642)	1.8122 *** (0.2607)	1.4276 *** (0.0775)	1.2828 *** (0.3992)
Escolaridad	0.0820 *** (0.0029)	0.0495 * (0.0302)	0.0834 *** (0.0030)	0.0783 ** (0.0370)
Experiencia	0.0229 *** (0.0025)	0.0286 *** (0.0052)	0.0501 *** (0.0131)	0.0965 *** (0.0352)
Experiencia ²	-0.0002 *** (0.0001)	-0.0004 *** (0.0001)	-0.0018 * (0.0010)	-0.0052 * (0.0027)
Experiencia ³			2.75E-05 (2.84E-05)	1.23E-04 (7.54E-05)
Experiencia ⁴			-1.26E-07 (2.72E-07)	-1.03E-06 (6.88E-07)
P1BE per cápita	0.0046 *** (0.0006)	0.0050 *** (0.0007)	0.0046 *** (0.0006)	0.0048 *** (0.0007)
N= 6488				

Cuadro 3: Estimación del modelo lineal, mínimos cuadrados ordinarios y variables instrumentales (continuación)

Región/Variable	1		11	
	MCO	V1	MCO	V1
Centro				
Constante	1.3019 *** (0.0721)	0.9677 *** (0.3043)	1.2201 *** (0.0887)	1.0944 *** (0.4202)
Escolaridad	0.0805 *** (0.0028)	0.1245 *** (0.0409)	0.0811 *** (0.0028)	0.1106 ** (0.0468)
Experiencia	0.0300 *** (0.0026)	0.0266 *** (0.0056)	0.0475 *** (0.0133)	0.0256 (0.0348)
Experiencia ²	-0.0004 *** (0.0001)	-0.0002 (0.0002)	-0.0015 (0.0010)	-0.0004 (0.0025)
Experiencia ³			2.52E-05 (2.98E-05)	1.06E-05 (6.91E-05)
Experiencia ⁴			-1.87E-07 (2.91E-07)	-1.54E-07 (6.30E-07)
P1BE per cápita	0.0094 *** (0.0009)	0.0088 (0.0012)	0.0094 *** (0.0009)	0.0090 *** (0.0011)
N= 11588				
Capital				
Constante	1.6359 *** (0.0682)	1.0096 ** (0.4672)	1.5773 *** (0.1090)	1.2589 *** (0.4113)
Escolaridad	0.0968 *** (0.0044)	0.1719 *** (0.0589)	0.0975 *** (0.0044)	0.1871 *** (0.0609)
Experiencia	0.0222 *** (0.0040)	0.0199 ** (0.0082)	0.0312 (0.0209)	-0.0807 (0.0657)
Experiencia ²	-0.0002 ** (0.0001)	-0.0001 (0.0002)	-0.0005 (0.0015)	0.0078 (0.0048)
Experiencia ³			-5.55E-07 (4.26E-05)	-2.31E-04 * (0.0001)
Experiencia ⁴			8.76E-08 (4.02E-07)	2.25E-06 * (1.26E-06)
P1BE per cápita	0.0007 *** (0.0002)	-0.0005 (0.0010)	0.0007 *** (0.0002)	-0.0008 (0.0010)
N= 2904				

Cuadro 3: Estimación del modelo lineal, mínimos cuadrados ordinarios y variables instrumentales (continuación)

Región/Variable	1		11	
	MCO	V1	MCO	V1
Sureste				
Constante	1.1807 *** (0.0589)	0.9454 ** (0.4531)	0.9800 *** (0.0854)	0.6892 * (0.3705)
Escolaridad	0.1010 *** (0.0032)	0.1319 ** (0.0557)	0.1022 *** (0.0032)	0.1483 *** (0.0449)
Experiencia	0.0367 *** (0.0033)	0.0318 *** (0.0072)	0.0783 *** (0.0161)	0.0615 (0.0404)
Experiencia ²	-0.0004 *** (0.0001)	-0.0003 (0.0002)	-0.0029 ** (0.0012)	-0.0025 (0.0030)
Experiencia ³			4.90E-05 (3.47E-05)	6.20E-05 (8.54E-05)
Experiencia ⁴			-2.92E-07 (3.35E-07)	-5.59E-07 (8.16E-07)
P1BE per cápita	0.0031 *** (0.0005)	0.0028 *** (0.0008)	0.0031 *** (0.0005)	0.0026 *** (0.0007)
N= 4856				
Sur				
Constante	0.7376 *** (0.1397)	0.7608 *** (0.2212)	0.4895 *** (0.1628)	0.2877 (0.3199)
Escolaridad	0.0986 *** (0.0033)	0.1931 *** (0.0694)	0.1002 *** (0.0034)	0.1492 *** (0.0559)
Experiencia	0.0321 *** (0.0037)	0.0006 (0.0165)	0.0953 *** (0.0194)	0.1470 ** (0.0705)
Experiencia ²	-0.0003 *** (0.0001)	0.0005 (0.0004)	-0.0049 *** (0.0015)	-0.0102 ** (0.0051)
Experiencia ³			0.0001 *** (4.27E-05)	0.0003 ** (0.0001)
Experiencia ⁴			-1.09E-06 *** (4.12E-07)	-2.78E-06 ** (1.29E-06)
P1BE per cápita	0.0186 *** (0.0043)	-0.0005 (0.0143)	0.0186 *** (0.0043)	0.0086 (0.0120)
N= 3518				

* Significativo al 90%; ** significativo al 95%; *** significativo al 99%; sin asterisco: no significativo

Fuente: Cálculos propios con información de la ENOE, tercer trimestre de 2005.

La frontera es la única región para la cual el impacto de la educación al usar VI es no significativo en la primera especificación, mientras que para el segundo modelo el coeficiente de escolaridad resultó estadísticamente significativo para todas las regiones. Con excepción del norte en las dos especificaciones, controlar por la endogeneidad de la educación genera retornos más altos.

Muchos estudios que han utilizado el método de VI han encontrado estimadores más altos con respecto a los obtenidos con MCO. Card (1999) realiza una revisión de algunos estudios empíricos realizados para Estados Unidos y otros países, y en los cuales se obtienen tales resultados; el autor otorga algunas explicaciones por las cuales esto puede estar ocurriendo. Card (1999) menciona que los estimadores VI pueden estar aún más sesgados a la alza que los correspondientes de MCO por diferencias no observadas entre las características de los grupos de tratamiento y de comparación implícitos en el método de VI. Esta explicación es más aceptable para estudios que no controlan por características familiares, como la presente investigación.

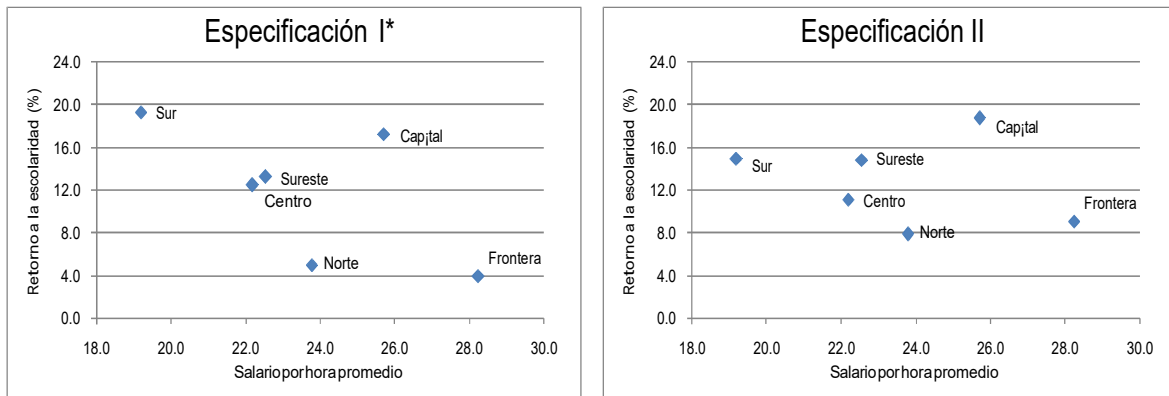
El autor también señala que los sesgos de habilidad presentes en los estimadores MCO de los retornos a la educación son relativamente pequeños, y que la diferencia entre los estimadores de VI y MCO puede estar reflejando el sesgo a la baja del estimador MCO atribuible a errores de medición. No obstante, el autor comenta que resulta extraño que grandes diferencias entre los estimadores VI y MCO se deban sólo a errores de medición.

Otra explicación puede deberse a la heterogeneidad en los retornos. Card (1999) afirma que factores como la obligatoriedad en la educación o el acceso a las escuelas es más probable que afecten las decisiones de los individuos que de otra manera habrían tenido baja escolaridad. Si esos individuos tienen retornos a la escolaridad más altos que el

promedio, entonces se esperaría que los estimadores VI basados en la obligatoriedad de la escolaridad o en la proximidad de las escuelas generaran retornos a la escolaridad más altos que MCO. Una condición necesaria para que esto ocurra es que los retornos a la escolaridad estén negativamente correlacionados con el nivel de escolaridad a través de la población, como ocurre en el presente estudio.

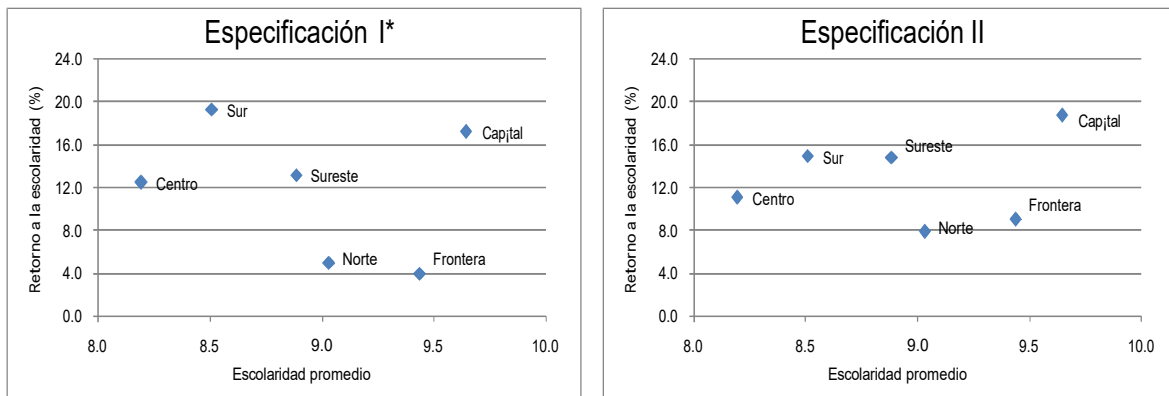
En las gráficas 5 y 6 se relaciona el retorno a la escolaridad con el ingreso y la escolaridad promedio, para las dos especificaciones. Si se excluye a la capital, la relación parece negativa; es decir, las regiones con menores niveles de desarrollo son las que alcanzan mayores tasas de retorno a la escolaridad. Sin embargo no se puede asegurar que tal relación sea la verdadera. Habría que buscar otros instrumentos que permitan confirmar que, al controlar por el carácter endógeno de la escolaridad, el desarrollo está inversamente relacionado con los rendimientos educacionales de las regiones mexicanas.

Gráfica 5: Relación de ingreso y retorno a la escolaridad (IV)



*El retorno de la frontera es no significativo

Gráfica 6: Relación de años de educación y retorno a la escolaridad



*El retorno de la frontera es no significativo

CONCLUSIONES

Los resultados encontrados en este trabajo son evidencia de rendimientos a la escolaridad diferenciados por región. Cuando los retornos educacionales son calculados por MCO se encuentra que estos difieren estadísticamente entre regiones.

Al comparar tales rendimientos con la escolaridad o el salario por hora promedios se observa cierta relación negativa: las regiones menos desarrolladas (como la sur y la sureste) son las que obtienen retornos más altos, y viceversa. Esto es acorde con la predicción neoclásica de convergencia del capital, y confirma los resultados de otros estudios realizados para México y otros países (Barceinas, 2005; Raymond, 2002). Lo anterior estaría indicando que destinar gasto público para la educación hacia las regiones más marginadas es óptimo desde el punto de vista de equidad y eficiencia. Con ello no solamente se estaría buscando incrementar el bienestar de tales regiones, sino que además el retorno de la inversión en educación sería más alto.

No obstante, debe considerarse que las estimaciones permiten obtener el *retorno marginal* de los años de escolaridad. En este sentido, sería útil obtener los rendimientos por niveles de educación (por región); de esta forma se podría distinguir a aquellos que presentan un mayor retorno, y por ende, hacia cuales es más rentable destinar recursos públicos. El ejercicio además permitiría entender mejor el proceso migratorio interregional; por ejemplo, si se confirmara que en las zonas más marginadas son los niveles altos de educación los que otorgan un menor rendimiento, esto explicaría por qué los individuos más calificados emigran hacia las regiones más desarrolladas.

Posteriormente y con el objetivo enfrentar el problema de endogeneidad de la educación, se utilizaron VI. En general se observa que, con excepción de la región norte,

considerar el carácter endógeno de la escolaridad incrementa los retornos a la educación; en particular, la capital muestra el rendimiento a la escolaridad más alto.

Si se relacionan los retornos a la educación con el ingreso o la escolaridad promedios, y si se excluye a la capital por considerarse un caso especial, se vuelve a encontrar cierta tendencia negativa (aunque no tan marcada como en el caso de MCO); es decir, las regiones más marginadas obtienen los rendimientos más altos. No obstante tal relación debe ser confirmada buscando otros instrumentos, debido a que no en todas las regiones se obtuvieron retornos a la educación estadísticamente significativos; además se deben de tener en mente los problemas que puede estar generando no tener información relevante (como los antecedentes familiares u otras variables) que determina la escolaridad y la experiencia individual.

Anexo 1: Regiones

Frontera	Baja California Chihuahua Coahuila Nuevo León Sonora Tamaulipas
Norte	Aguascalientes Baja California Sur Durango Nayarit San Luis Potosí Sinaloa
Centro	Colima Guanajuato Hidalgo Jalisco Michoacán Morelos Puebla Querétaro Tlaxcala Veracruz Zacatecas
Capital	Distrito Federal México
Sureste	Campeche Quintana Roo Tabasco Yucatán
Sur	Chiapas Guerrero Oaxaca

Anexo 2. Intervalos de confianza para las medias de las variables, al 95%

Región	Salario por hora		Edad		Escolaridad		Experiencia	
Frontera	28.226	28.249	35.094	35.103	9.435	9.439	19.657	19.666
Norte	23.765	23.789	35.429	35.438	9.029	9.035	20.395	20.407
Centro	22.178	22.193	35.456	35.462	8.192	8.196	21.261	21.269
Capital	25.682	25.734	35.521	35.539	9.640	9.649	19.876	19.896
Sureste	22.524	22.560	35.335	35.346	8.880	8.888	20.450	20.463
Sur	19.168	19.202	35.025	35.043	8.502	8.515	20.515	20.537

Anexo 3. Intervalos de confianza para los efectos marginales de la educación

Región	Intervalo de confianza al 95%	
Frontera	7.39381	7.40419
Norte	7.24721	7.25871
Centro	8.04395	8.05483
Capital	7.57353	7.58752
Sureste	9.30959	9.32045
Sur	9.84791	9.86191

BIBLIOGRAFÍA

- Barceinas, F., *"Endogeneidad y rendimientos de la educación"*, Estudios Económicos, vol. 18, núm. 1, Enero-Junio de 2003, págs. 79-131.
- Barceinas, F., *"Convergencia regional y capital humano en México, de los años 80 al 2002"*, Estudios Económicos, vol. 20, no. 2, Julio-Diciembre 2005, págs. 263-302.
- Bracho, T., Zamudio, A., *"Los rendimientos económicos de la escolaridad en México"*, 1989, Economía Mexicana. Nueva Época, vol. III, núm. 2, segundo semestre de 1994.
- Cahuc, P., Zylberberg, A. *"Labor Economics"*, MIT Press, 2004.
- Cameron, A. C., Trivedi, P. K., *"Microeconometrics: methods and applications"*, Cambridge University Press, 2005.
- Card, D., *"The Causal Effect of Education on Earnings"*, Handbook of Labor Economics, Volume 3, Edited by O. Ashenfelter and D. Card, 1999.
- Guizar, J., *"Los retornos a la educación y la productividad del trabajo en México". ¿Hay un efecto sheepskin?*, Tesis de Licenciatura, CIDE, Junio 2007.
- Hanson, G., Woodruff, C., *"Emigration and Educational Attainment in Mexico"*, University of California at San Diego, Mimeographed, 2003.
- Button, K.J., *"Infrastructure investment, endogenous growth and economic convergence"*, The Annals of Regional Science, Vol. 32, 1, pp. 145-162.
- Lemieux, T., *"The 'Mincer Equation' Thirty Years after Schooling, Experience, and Earnings"*, Jacob Mincer: A Pioneer of Modern Labor Economics, Edited by Shoshana Grossbard, New York, Springer, 2006.
- López-Acevedo, G., *"Evolution of Earnings and Rates of Returns to Education in Mexico"*, Latin America and the Caribbean Region, Poverty Reduction and Economic

Management Division, The World Bank, 2001.

Messmacher-Linartas, M., *"Desigualdad regional en México. El efecto del TLCAN y otras reformas estructurales."*, Documento de Investigación No. 2000-4, Dirección General de Investigación Económica, Banco de México.

Mincer, J., *"Schooling, Experience and Earnings"*, USA, NBER, New York, 1974.

-----*"Changes in Wage Inequality, 1970-1990"*, Research in Labor Economics, 16, 1-18, 1997.

Raymond, J.L., *"Convergencia real de las regiones españolas y capital humano"*, Papeles de Economía Española, No. 93, 2002. Convergencia regional: España y Europa.

Wooldridge, J.M., *"Econometric Analysis of Cross Section and Panel Data"*, MIT Press, 2002.

Zamudio, A., *"Rendimientos a la educación superior en México: ajuste por sesgo utilizando máxima verosimilitud"*, Economía Mexicana. Nueva Época, vol. IV, núm. 1, primer semestre de 1995.