

# ESTUDIOS DE LA FUNDACIÓN

SERIE ECONOMÍA Y SOCIEDAD

## ■ ¿ES RENTABLE EDUCARSE?

**MARCO CONCEPTUAL Y PRINCIPALES  
EXPERIENCIAS EN LOS CONTEXTOS  
ESPAÑOL, EUROPEO Y EN PAÍSES  
EMERGENTES**

---

José Luis Raymond  
(Coordinador)



FUNDACIÓN DE LAS CAJAS DE AHORROS





ESTUDIOS  
DE LA FUNDACIÓN

SERIE ECONOMÍA Y SOCIEDAD





**ESTUDIOS**  
DE LA FUNDACIÓN

SERIE **ECONOMÍA Y SOCIEDAD**



**¿ES RENTABLE  
EDUCARSE?**

**MARCO CONCEPTUAL Y PRINCIPALES  
EXPERIENCIAS EN LOS CONTEXTOS  
ESPAÑOL, EUROPEO Y EN PAÍSES  
EMERGENTES**

---

José Luis Raymond  
(Coordinador)



FUNDACIÓN DE LAS CAJAS DE AHORROS

## FUNDACIÓN DE LAS CAJAS DE AHORROS

## PATRONATO

ISIDRO FAINÉ CASAS (*Presidente*)  
JOSÉ MARÍA MÉNDEZ ÁLVAREZ-CEDRÓN (*Vicepresidente*)  
FERNANDO CONLLEDO LANTERO (*Secretario*)  
JULIO FERNÁNDEZ GAYOSO  
MARIO FERNÁNDEZ PELAZ  
JORDI MESTRE GONZÁLEZ  
ANTONIO PULIDO GUTIÉRREZ  
ATILANO SOTO RÁBANOS  
ADOLFO TODO RÓVIRA

## DIRECTOR GENERAL

VICTORIO VALLE SÁNCHEZ

Printed in Spain

Edita: FUNDACIÓN DE LAS CAJAS DE AHORROS (FUNCAS)  
Caballero de Gracia, 28, 28013 - Madrid

© FUNDACIÓN DE LAS CAJAS DE AHORROS (FUNCAS)

Todos los derechos reservados. Queda prohibida la reproducción total o parcial de esta publicación, así como la edición de su contenido por medio de cualquier proceso reprográfico o fónico, electrónico o mecánico, especialmente imprenta, fotocopia, microfilm, *offset* o mimeógrafo, sin la previa autorización escrita del editor.

ISBN: 978-84-89116-70-2

Depósito legal: M. 24016-2011

Preimpresión: Advantia, S.A.

Imprime: Advantia, S.A.

PRESENTACIÓN	11
CAPÍTULO 1. EL EFECTO CAUSAL DE LA EDUCACIÓN SOBRE LOS SALARIOS	17
<i>David Card</i>	
RESUMEN	19
1.1. INTRODUCCIÓN	19
1.2. LA FUNCIÓN DE INGRESOS DEL CAPITAL HUMANO	21
1.2.1. Forma funcional	22
1.2.2. Medición de la educación	24
1.2.3. ¿Qué medida de ingresos?	26
1.2.4. Resumen	27
1.3. MODELIZACIÓN CAUSAL DEL RENDIMIENTO DE LA EDUCACIÓN	28
1.3.1. Aspectos teóricos	28
1.3.2. Educación observada e ingresos obtenidos	32
1.3.3. Errores de medida	35
1.3.4. Estimaciones con variables instrumentales del rendimiento de la educación	36
1.3.5. Limitaciones de las variables instrumentales	38
1.3.6. <i>Background</i> familiar	42
1.3.7. Modelos para hermanos y gemelos	46
1.3.8. Resumen	52
1.4. UNA REVISIÓN SELECTIVA DE ESTUDIOS EMPÍRICOS RECIENTES	54
1.4.1. Variables instrumentales basadas en aspectos institucionales del sistema educativo	54
1.4.2. Estimaciones empleando <i>background</i> familiar como control o como instrumento	63
1.4.3. Estudios de educación e ingresos con gemelos	67
1.4.4. Evidencia directa de la heterogeneidad en los rendimientos educativos	72
1.5. CONCLUSIONES	76
APÉNDICE	78
BIBLIOGRAFÍA	81

<b>CAPÍTULO 2. LOS RENDIMIENTOS DE LA INVERSIÓN EN EDUCACIÓN: UNA ACTUALIZACIÓN ADICIONAL</b>	<b>87</b>
<i>George Psacharopoulos y Harry Anthony Patrinos</i>	
RESUMEN	89
2.1. INTRODUCCIÓN	89
2.2. PATRONES RECIENTES	90
2.3. UN ENFOQUE MÁS SELECTIVO	95
2.4. EXTENSIONES	97
2.5. ASPECTOS DE POLÍTICA	98
2.6. CONCLUSIONES	99
2.7. BIBLIOGRAFÍA	100
APÉNDICE	105
<b>CAPÍTULO 3. LA RENTABILIDAD DE LA EDUCACIÓN: ENFOQUE MICROECONÓMICO</b>	<b>119</b>
<i>Colm Harmon, Hessel Oosterbeek e Ian Walter</i>	
RESUMEN	121
3.1. INTRODUCCIÓN	121
3.2. EL ENFOQUE DE CAPITAL HUMANO Y LA RENTABILIDAD DE LA EDUCACIÓN	122
3.2.1. La especificación de Mincer	122
3.2.2. Elección óptima de escolaridad	123
3.2.3. Sesgo de habilidad	126
3.3. ANÁLISIS DE REGRESIÓN	126
3.3.1. Análisis comparativo	126
3.3.2. Especificación y forma funcional	128
3.3.5. Medidas alternativas de la educación alcanzada	134
3.3.6. Variación de los rendimientos educativos a través de la distribución salarial	135
3.3.7. Metanálisis	137
3.3.8. Otras fuentes de variación de los rendimientos: Sobreeducación	139
3.4. SEÑALIZACIÓN	141

3.5. ENDOGENEIDAD DE LA EDUCACIÓN	144
3.5.1. Aislando el efecto de variación exógena en la escolaridad	144
3.5.2. Resultados de los Estudios con Variables Instrumentales –Evidencia Internacional	147
3.5.3. ¿Por qué son las estimaciones por VI mayores a las MCO?	153
3.5.4. Relevancia del instrumento y validez del instrumento:	154
3.5.5. Evidencia adicional–estimadores de efectos fijos y métodos de correspondencia	155
3.6. CONCLUSIONES	157
BIBLIOGRAFÍA	160
<b>CAPÍTULO 4. CAPITAL HUMANO Y RENDIMIENTOS DE LA EDUCACIÓN EN ESPAÑA</b>	<b>167</b>
<i>José L. Raymond, José L. Roig, Gustavo A. García y Lina M. Gómez</i>	
4.1. INTRODUCCIÓN	169
4.2. CAPITAL HUMANO Y RENDIMIENTOS EDUCATIVOS EN ESPAÑA. ALGUNOS ESTUDIOS DE SU EVOLUCIÓN Y DISTRIBUCIÓN GEOGRÁFICA	171
4.3. BASE DE DATOS DISPONIBLES Y HOMOGENEIZACIÓN. ANÁLISIS DESCRIPTIVO	172
4.4. ANÁLISIS DEL COMPORTAMIENTO DEL CAPITAL HUMANO POR REGIONES Y SECTORES	179
4.5. EVOLUCIÓN DE LOS RENDIMIENTOS EDUCATIVOS EN EL PERÍODO 1995, 2002 Y 2006: DIFERENCIAS POR SEXO Y CC.AA.	184
4.5.1. Comportamiento de los rendimientos educativos según sexo	188
4.5.2. Los retornos educativos por comunidad autónoma	190
4.5.3. Factores condicionantes de las diferencias salariales para España en la última década	191
4.6. LA PROBABILIDAD DE EMPLEO Y LOS RENDIMIENTOS EDUCATIVOS	197
4.6.1. La consideración de las probabilidades de empleo en la evaluación de los rendimientos educativos	197
4.6.2. Efecto de la probabilidad de empleo sobre la estimación de los retornos en España	202
4.7. EL SESGO DE HABILIDAD Y LA FORMACIÓN DE UN PSEUDO PANEL	208
4.8. EL <i>BACKGROUND</i> FAMILIAR Y SESGO DE ENDOGENEIDAD: APLICACIÓN DE VARIABLES INSTRUMENTALES	217

4.9. DETERMINANTES DE LOS RENDIMIENTOS EDUCATIVOS	222
4.9.1. Los retornos educativos y la relación con el paro	227
4.9.2. Efecto del desempleo sobre los rendimientos educativos corregidos por la probabilidad de ocupación	228
4.10. EXTERNALIDADES DEL CAPITAL HUMANO	231
4.10.1. Función de producción y salarios	231
4.10.2. Estimaciones de las externalidades de la acumulación de capital humano	232
4.11. CONCLUSIONES	235
REFERENCIAS	237
ANEXOS ESTADÍSTICOS	239



## **PRESENTACIÓN**



## PRESENTACIÓN

El bienestar económico de una sociedad, entre otros factores, depende de la cantidad de bienes y servicios que en términos per cápita esta sociedad sea capaz de producir. Y esta capacidad productiva se expresa a través de la conocida función de producción que liga el output con los *inputs* fuerza de trabajo, capital humano o calidad de esta fuerza de trabajo, capital físico y recursos naturales. De estos cuatro *inputs* contemplados, el capital humano en sentido amplio posiblemente sea el más relevante como la experiencia reciente prueba. Muchos países quedaron destruidos durante la segunda guerra mundial pero aquellos que tenían un buen *stock* de capital humano, como Alemania o Japón, pronto recuperaron posiciones a pesar de que su capital físico había quedado destruido. Por otro lado, la experiencia de muchos países bien dotados de recursos naturales pero que, a pesar de ello, no logran despegar, mientras que países mal dotados de recursos ocupan las primeras posiciones en el *ranking* del desarrollo económico, es una prueba de que la abundancia de recursos naturales no es condición suficiente ni necesaria para el desarrollo. El despegue económico de los países se ha producido siempre espoleado por una mejora de la productividad y esta mejora de la productividad se ha sustentado en el cambio técnico en sentido amplio. El cambio técnico exige innovar, o exige capacidad para imitar con éxito pautas productivas y organizativas de países más desarrollados del entorno. Y sin duda, innovar, o imitar con éxito pautas productivas y organizativas, hacen que el capital humano se erija como un requisito indispensable. La forma usual de medir este capital humano es a través del nivel educativo de la población en edad de trabajar y analizando también la contribución de este capital humano al proceso de generación de renta.

Conscientes de la importancia del tema, los autores hemos llevado a cabo una investigación acerca de los rendimientos de la educación en la economía española. Pero en el momento de publicar los resultados obtenidos nos ha parecido oportuno ubicar esta contribución en el contexto de una extensa literatura que se ha ocupado del tema.

El libro consta de cuatro contribuciones<sup>1</sup>. La primera de ellas ofrece el marco conceptual general que sirve de base para la medición de los rendimientos del capital humano. La

---

<sup>1</sup> La traducción de los capítulos 1 a 3 ha sido realizada por Lina Gómez, en tanto que su revisión técnica ha sido efectuada por José Luis Raymond y José Luis Roig. Por motivos editoriales, en la estructura de esos tres capítulos ha sido respetado el formato en el que fueron originalmente publicados.

segunda presenta los resultados obtenidos en el estudio de los rendimientos de la educación en un muy amplio espectro de países. Podríamos referir esta contribución como el análisis de los rendimientos de la educación en el mundo. La tercera contribución analiza los rendimientos de la educación en Europa. Y, finalmente, la cuarta contribución presenta los rendimientos de la educación en España y sus CCAA.

El primer trabajo seleccionado es el de David Card titulado “El efecto causal de la educación sobre los salarios” e inicialmente publicado en el *Handbook of Labor Economics* (1999) si bien se trata de un artículo que puede que presente alguna dosis de dificultad, nos ha parecido que ofrece una excelente aproximación conceptual a las cuestiones que plantea evaluar la contribución de la educación al proceso de generación de renta. A pesar de no ser una contribución reciente, pues data de 1999, nos ha parecido que ofrece la aproximación más completa al tema tanto por su extensión como por sus múltiples referencias a las cuestiones complejas que plantea mediar el efecto causal de una variable en un contexto en el que, en general, no es factible el desarrollo de experimentos controlados.

El trabajo de Psacharopoulos y Patrinos titulado “Los rendimientos de la inversión en educación: Una actualización adicional” fue publicado en 2004 en *Education Economics*. Es un estudio de menores aspiraciones conceptuales pero ofrece la más completa comparación de los rendimientos educativos para un amplio espectro de países y utilizando especificaciones econométricas y bases de datos que tratan de garantizar su comparabilidad. Como es bien conocido, un modelo de regresión es una forma de realizar inferencia condicionada. Si en una ecuación de salarios no se mantiene la misma especificación y el mismo método de estimación, los coeficientes de regresión estimados dejan de ser comparables. La insistencia en estos extremos consideramos que es el principal valor añadido que el trabajo de Psacharopoulos y Patrinos ofrece.

Los rendimientos de la educación en Europa se presentan en el trabajo de Harmon, Oosterbeek y Walker titulado “Los rendimientos de la educación: Un enfoque microeconómico” e inicialmente publicado en 2003 en el *Journal of Economic Surveys*. Aunque la palabra “Europa” no aparezca en el título, el artículo fundamentalmente revisa los rendimientos educativos en los países europeos aprovechando la información que dos de sus autores (Harmon y Walker) tuvieron como participantes en el proyecto “PURE” (*Public Funding and Private Returns to Schooling*). Este proyecto, patrocinado por la Comunidad Europea, tuvo por objeto examinar los rendimientos privados de la educación en los principales países europeos, incluida España, y se puso especial cuidado en tratar de conseguir también bases de datos y especificaciones econométricas comparables.

La última contribución es la de Raymond, Roig, García y Gómez que examina los rendimientos de la educación en España y en sus CCAA. Se trata de una investigación patrocinada por FUNCAS en la que, básicamente, se explota la información que ofrece la Encuesta sobre Estructura Salarial que elabora el INE y que permite una comparación homogénea de los rendimientos educativos en los años 1995, 2002 y 2006. Adicionalmente, el tamaño

de la muestra es tan elevado que es factible la separación por CCAA. Se constata que a pesar del fuerte incremento que la oferta de educados ha tenido en España, educarse sigue siendo rentable, y que esta rentabilidad procede de dos mecanismos. Los más educados perciben una prima salarial, y los más educados consiguen una mejor inserción en el mercado laboral, a pesar de los problemas de sobreeducación que en algunos casos puedan darse.

**Los autores**

Febrero de 2011





# 1

## EL EFECTO CAUSAL DE LA EDUCACIÓN SOBRE LOS SALARIOS

David Card (2001)\*

---

\*Agradezco a David Lee y Gena Estes por la asistencia en la investigación, a Orley Ashenfeller, Alan Krueger y James Powell por los útiles debates, y a Michael Boozer, Ken Chay, Andrew Hildreth y Gary Solón por los comentarios sobre borradores preliminares que mejoraron sustancialmente este capítulo. Esta investigación fue financiada en parte por una beca del NICHD.



# 1. EL EFECTO CAUSAL DE LA EDUCACIÓN SOBRE LOS SALARIOS\*

## ■ RESUMEN

Este documento revisa la literatura reciente sobre la relación causal entre educación e ingresos. Se centra en cuatro áreas de trabajo: avances teóricos y econométricos en la modelación del efecto causal de la educación en presencia de heterogeneidad en los rendimientos educativos; estudios recientes que utilizan aspectos institucionales del sistema educativo para obtener la estimación del rendimiento de la educación por variables instrumentales; estudios recientes de la relación entre ingresos y educación en gemelos; y recientes intentos para evidenciar las fuentes de heterogeneidad en los rendimientos educativos. Consistentes con las primeras revisiones de literatura, se concluye que en promedio (o promedio marginal) el rendimiento educativo no es muy inferior a las estimaciones que resultan de una función estándar de ingresos de capital humano ajustada por MCO. La evidencia de los últimos estudios con gemelos idénticos sugiere un pequeño sesgo de “habilidad” al alza —del orden del 10%. Un resultado consistente entre estudios que utilizan variables instrumentales basadas en cambios institucionales en el sistema educativo, es que los rendimientos estimados son un 20-40% superiores a las correspondientes estimaciones MCO. Parte de la explicación de este resultado puede estar en que la rentabilidad marginal de la educación para ciertos grupos —particularmente grupos relativamente desaventajados con bajos niveles educativos— es mayor que la rentabilidad marginal promedio de la educación para el conjunto de la población.

*Códigos JEL: I20, J30.*

## ■ 1.1 INTRODUCCIÓN

La educación juega un papel central en los mercados laborales modernos. Cientos de estudios, para países y periodos de tiempo diferentes, confirman que los individuos mejor educados ganan salarios más altos, experimentan menor desempleo, y trabajan en ocupaciones más pres-

---

\* David Card, “The Causal Effect of Education on Earnings”, capítulo 30 del *Handbook of Labour Economics*, volumen 3, editado por O. Ashenfelter y D. Card, 1999, Elsevier Science, páginas 1801-1863.

tigiosas que sus homólogos menos educados<sup>1</sup>. A pesar de la contundente evidencia sobre la correlación positiva entre educación y estatus en el mercado laboral, los científicos sociales han sido muy prudentes al no derivar inferencias fuertes sobre el efecto causal de la educación. En ausencia de evidencia experimental, es muy difícil saber si los mayores ingresos observados para los trabajadores mejor educados son causados por su mayor educación, o si individuos con mayor capacidad de generación de ingresos han elegido adquirir mayor educación.

El interés de los economistas sobre este problema fue estimulado a finales de los cincuenta por los ejercicios de contabilidad del crecimiento que encontraron que los crecientes niveles de educación podían explicar gran parte del crecimiento de la productividad de EE.UU. en la postguerra, relegando el cambio técnico a un papel menor (véase, por ejemplo, Becker, 1964; Griliches, 1970). Los escépticos argumentaban que esta conclusión solo era válida si las diferencias de ingresos observadas de ingresos en corte transversal entre grupos educativos reflejasen diferenciales de productividad genuinos, y no diferencias de habilidad inherente que estuviesen correlacionadas con la educación (p. ej., Deninson, 1964). La aparición de bases de datos microeconómicas de gran tamaño muestral en los años sesenta, llevó a la generación de una gran cantidad de investigación sobre educación y salarios, mucha de ella centrada en el problema del “sesgo de habilidad” en los diferenciales salariales entre los trabajadores más y menos educados. En su conocida revisión de la literatura de los años sesenta y setenta, Griliches (1977) concluyó que tales sesgos eran pequeños—incluso potencialmente menores que otros sesgos que llevan a las diferencias de salarios estimadas subestimar el efecto causal de la educación. En su revisión más temprana de la evidencia empírica, Becker (1964) igualmente había concluido que los sesgos habilidad eran sobreestimados por los críticos del paradigma de capital humano<sup>2</sup>. A pesar del riguroso razonamiento de estos primeros *surveys*, no obstante, muchos analistas continúan creyendo que la correlación parcial medida entre educación e ingresos sobreestima significativamente el verdadero efecto causal de la educación, y que resultados opuestos son contraintuitivos.

El propósito de este capítulo es revisar e interpretar la evidencia más reciente sobre la relación causal entre educación e ingresos. Centro mi atención en cuatro áreas fundamentales de investigación:

- 1) avances teóricos y econométricos en la modelización del efecto causal de la educación en la presencia de heterogeneidad en la rentabilidad de la educación;
- 2) estudios recientes que usan aspectos institucionales del sistema educativo como fuente “exógena” de variación en los niveles educativos;

<sup>1</sup> Véase Cohn y Addison (1997) para una revisión selectiva de estudios internacionales recientes, y Psacharopoulos (1985, 1994) para una extensa visión general de la literatura sobre educación e ingresos.

<sup>2</sup> Becker (1964, pp. 88, n. 30) ofreció la siguiente interpretación de la opinión predominante sobre la importancia de los sesgos de habilidad: “una explicación más cínica sería que los investigadores son fundamentalmente exitosos graduados de universidades y, por lo tanto, naturalmente, sesgados hacia la opinión de que la habilidad es una de las principales causas de los altos ingresos que reciben los graduados universitarios”.

- 3) estudios recientes sobre salarios y resultados de la educación en gemelos;
- 4) estudios recientes que modelan explícitamente la heterogeneidad en el rendimiento de la educación entre grupos o individuos.

Un tema unificador en gran parte de este trabajo es que el rendimiento de la educación no es un parámetro único en la población, sino una variable aleatoria que puede variar con otras características de los individuos, como el *background* familiar, la habilidad o el nivel de educación. En mi opinión, esta visión más amplia del efecto de la educación ayuda a conciliar los diversos hallazgos de la literatura, y proporciona un marco útil para generar nuevas hipótesis y perspectivas sobre la relación entre educación e ingresos.

El capítulo comienza con una breve descripción de la llamada función de ingresos de capital humano, que es el modelo básico que utilizan los economistas para medir el rendimiento educativo. Seguidamente presento una discusión en profundidad de un sencillo modelo de donde se endogeneiza la educación, que es útil para interpretar recientes estudios empíricos. Finalmente, presento una selectiva revisión y síntesis de algunos de los recientes trabajos más interesantes sobre educación e ingresos.

## ■ 1.2. LA FUNCIÓN DE INGRESOS DEL CAPITAL HUMANO

Los estudios recientes sobre educación y determinación salarial están casi siempre enmarcados en el enfoque de la función de ingresos del capital humano (FICH) de Mincer (1974). Según este modelo, el logaritmo del ingreso del individuo ( $y$ ) en un periodo de tiempo dado se puede descomponer en una función aditiva de un término lineal de la educación y un término cuadrático de la experiencia:

$$\log y = a + bS + cX + dX^2 + e \quad (1)$$

donde  $S$  representa los años de educación terminados,  $X$  el número de años que un individuo ha trabajado desde que terminó de estudiar, y  $e$  es un residuo estadístico. En ausencia de información directa sobre experiencia, Mincer propuso el uso de la "experiencia potencial": el número de años que podría haber trabajado un individuo con  $A$  años de edad, suponiendo que empezó a estudiar a los 6 años, terminó  $S$  años de educación en exactamente  $S$  años, y empezó a trabajar inmediatamente después:  $XA - S - 6$ . Aunque Mincer derivó esta ecuación de un modelo teórico de decisiones de elección educacional y de formación post-escolar, el patrón básico de variación de los ingresos por edad y educación se conoce al menos desde principios de los años cincuenta (p.ej., Miller, 1955)<sup>3</sup>. Así la FICH puede ser

<sup>3</sup> Miller (1955, pp 64-67) presenta los perfiles de edad de los datos de ingresos anuales para hombres en el Censo de 1950, para tres grupos diferentes de educación, y resalta la naturaleza cóncava de estos perfiles, y el hecho de que el perfil para los hombres con mayor nivel educativo obtiene el máximo 10 años después que el perfil de los hombres con menor nivel educativo. El análisis de Miller con los datos del Censo de la década de los sesenta confirmó estas mismas tendencias (Miller, 1966).

vista como un matrimonio extraordinariamente exitoso entre el razonamiento inductivo y deductivo.

### ■ 1.2.1. Forma funcional

La sencilla especificación de la ecuación 1, inmediatamente plantea una serie de preguntas que se han intentado responder directa e indirectamente en los últimos 20 años. Muchas de ellas respecto a la forma funcional. La ecuación de Mincer se puede considerar una aproximación a una forma funcional general,

$$\log y = F(S,A) + e$$

dado que  $S$  y  $A$  son medidas como variables discretas en la mayoría de bases de datos, la función  $F(\cdot)$  puede ser estimada no paramétricamente al incluir un conjunto completo de variables *dummy* para todos los pares  $(S,A)$ , o empleando métodos no paramétricos de suavización (por ejemplo, estimadores de kernel de densidad) en bases de datos más pequeñas<sup>4</sup>. Alternativamente, los investigadores han añadido términos de mayor orden en educación y edad o experiencia en la ecuación 1, y han examinado la mejora del ajuste respecto a la especificación original de Mincer. Un exhaustivo estudio en esta línea realizado por Murphy y Welch (1990), alcanzaba la conclusión de que una generalización del modelo de Mincer

$$\log y = a + bS + g(X) + e \quad (1')$$

donde  $g$  es un polinomio de tercer o posiblemente cuarto grado, produce una mejora significativa en el ajuste.

Evidencia reciente sobre la forma de la función  $F(\cdot)$  y el comportamiento de una especificación como la ecuación (1'), se presenta en el gráfico 1.1 que muestra el perfil de edad-ingreso para hombres y mujeres empleando el pool de las muestras de marzo de 1994, 1995 y 1996 del *Current Population Survey*. Los datos representan el ingreso por hora promedio para cada año de edad, para individuos con 10, 12 y 16 años de educación. Graficado junto con la media observada están los valores predichos de modelos como (1') que incluyen un término cúbico en la experiencia potencial<sup>5</sup>. Comparaciones entre datos obser-

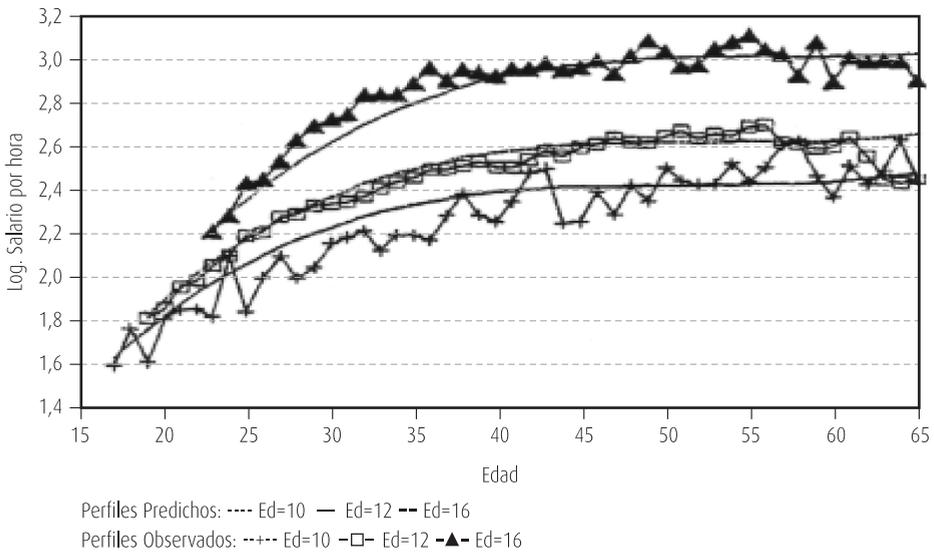
<sup>4</sup> En la mayoría de bases de datos estadounidenses, por ejemplo,  $S$  tiene valores discretos alrededor de 18 o 20, y  $A$  abarca desde los 16 hasta los 66 años de edad, lo que implica un máximo de aproximadamente 1000 puntos en el rango  $F(\cdot)$ . Zheng (1996) utiliza métodos formales de contrastación para comparar el ajuste de varias versiones ampliadas de (1), con las estimaciones del kernel de densidad, empleando datos del *Current Population Survey* para marzo de 1990.

<sup>5</sup> Las muestras incluyen 102.718 hombres y 95.360 mujeres, entre 16 a 66 años de edad, con experiencia potencial positiva e ingresos por hora promedio de entre \$ 2.00 y \$ 150.00 en dólares de 1995. Se presentan en los gráficos los individuos con 10, 12 o 16 años de educación (el cincuenta y tres por ciento de la muestra). Los modelos de regresión se ajustan por género para todos los grupos de educación, e incluyen un término lineal de educación, uno cúbico en la experiencia, y una variable *dummy* para individuos de raza negra.

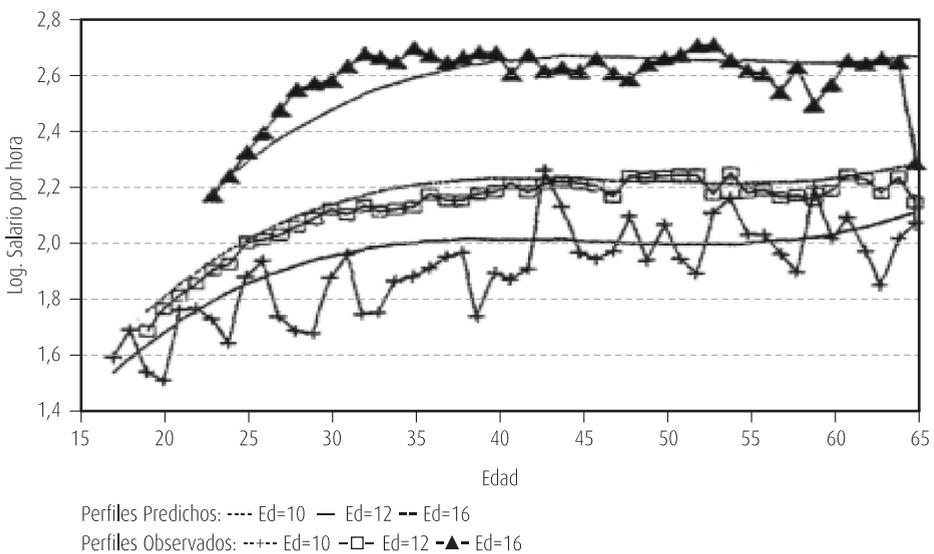
Gráfico 1.1

## PERFILES DE EDAD DEL SALARIO POR HORA PARA HOMBRES (a) Y MUJERES (b)

a. Perfiles de Salario por hora para hombres



b. Perfiles de Salario por hora para mujeres



vados y predichos sugieren que los perfiles de edad-ingreso para hombres y mujeres en EE.UU. son bastante suaves, y estarían razonablemente bien aproximados por una versión sencilla de la función de ingresos de capital humano estándar. No obstante, incluso una versión cúbica del modelo de Mincer tiene algunos problemas para adaptarse a la curvatura precisa de los perfiles de edad para diferentes grupos de educación, en datos de los últimos años para EE.UU. En particular, los modelos ajustados tienden a subestimar la tasa de crecimiento de los ingresos de los hombres y mujeres jóvenes con estudios universitarios respecto a los graduados de secundaria, lo que sugiere la necesidad de interacciones más flexibles entre educación y experiencia. Para algunos propósitos estos errores de especificación pueden no tener mucha importancia. En otras aplicaciones, no obstante, los sesgos en los perfiles de edad predichos de diferentes grupos de educación pueden llevar a graves confusiones.

### ■ 1.2.2. Medición de la educación

Además de imponer separabilidad entre los efectos de la educación y la experiencia, la función de ingresos de capital humano estándar supone que el logaritmo del ingreso es una función *lineal* de los años de educación terminados. Hay dos hipótesis implícitas (y relacionadas) en esta especificación: la primera, que la medida indicada para la educación es el número de años terminados de educación; y la segunda, que cada año adicional tiene el mismo efecto proporcional sobre los salarios, manteniendo constante los años en el mercado laboral. Suponiendo que estas condiciones son satisfechas, el coeficiente  $b$  de la ecuación (1) sintetiza completamente el efecto de la educación en el mercado laboral. Ahora es convencional referirse a  $b$  como “el rendimiento de la educación”<sup>6</sup>. Como se muestra en Willis (1986, p. 532) si (1) o (1') está correctamente especificadas, entonces  $b$  es de hecho la tasa interna de retorno de las inversiones en educación, suponiendo que la educación es gratuita y que los estudiantes no obtienen ningún ingreso mientras están en la escuela.

El uso de los años de educación terminados como medida de educación tiene una larga trayectoria en EE.UU. Estos datos fueron recolectados en los Censos Decenales de 1940-1980, y en los *Current Population Surveys* (CPS) desde 1940 hasta principios de 1990. La utilización de años de educación tiene validez sustancial en el sistema educativo de EE.UU., pero es menos natural en países con múltiples vías educativas (por ejemplo, Alemania o Francia) donde la graduación de escuela secundaria puede implicar diferentes años

<sup>6</sup> De hecho, el coeficiente de la educación en cualquier modelo estadístico de salarios (o ingresos) es generalmente considerado como el “rendimiento de la educación”, independientemente de qué otras variables de control son incluidas en el modelo. Esto puede generar alguna confusión cuando la edad, en lugar de la experiencia potencial ( $X$ ) se incluye como control, ya que la derivada de la ecuación (1) con respecto a la educación manteniendo la edad constante es  $b - c - 2dX$ . En consecuencia, el “rendimiento de la educación” es generalmente menor en los modelos que controlan por edad, en lugar de hacerlo por la experiencia (Mincer, 1974, p. 84).

de educación en función de si un estudiante piensa ir a la universidad, a la escuela de formación profesional, o empezar a trabajar de inmediato<sup>7</sup>.

Incluso entre muchos analistas estadounidenses se ha argumentado que las credenciales (como el diploma de secundaria o el grado universitario) importan más que los años de educación en sí mismos. Esta hipótesis ha sido conocida como “efecto credencialismo” —la existencia de primas salariales por el cumplimiento del último año de la escuela primaria, la escuela secundaria o la universidad. Hungerford y Solon (1987), y Belman y Heywood (1991) aumentan una función de ingresos estándar como la de la ecuación (1), con variables que capturan no linealidades a los 8, 12 o 16 años de educación. Estos autores encontraron evidencia de no linealidades, especialmente alrededor de los 16 años de educación (correspondiente al grado universitario)<sup>8</sup>. Park (1994) analizó una muestra grande de datos de la CPS y concluyó que gran parte de la aparente no linealidad a los 16 años de educación se debía a diferencias relativamente pequeñas entre el ingreso de los individuos con 14 y 15 años de educación (es decir, una rentabilidad excepcionalmente baja a los 15 años de educación, más que una excepcionalmente alta a los 16 años de educación). Adicionalmente, Park muestra que la forma funcional lineal produce un ajuste sorprendentemente bueno de los datos.

A pesar de la satisfacción general de los economistas con la medida tradicional de la educación, a finales de los años ochenta la Oficina Censal de EE.UU. decidió cambiar a un sistema de medición basado en grados de educación post secundaria (véase Kominski y Siegel, 1992). Así, a los individuos en el Censo de 1990 y los *Current Population Surveys* recientes, ya no se les preguntó cuántos años de universidad habían completado: se les preguntó sobre la posesión o no de título. Este cambio dificulta la estimación del modelo de ingresos de capital humano estándar con los datos recientes de EE.UU, o medir cambios en la estructura de las diferencias de salarios relacionadas con la educación. Sin embargo, se puede construir una concordancia entre la antigua variable años de educación y la nueva variable basada en los títulos a partir de una tabulación cruzada de las respuestas a las dos preguntas incluidas en un suplemento del CPS de febrero de 1990. El uso de esta concordancia apoya sorprendentemente el supuesto de linealidad implícito en la especificación original de Mincer<sup>9</sup>.

El gráfico 1.2 muestra los datos de salarios y educación para una muestra de hombres entre 40-55 años en el CPS de 1994-1996<sup>10</sup>. La media del logaritmo de los salarios para cada grupo educativo (por ejemplo, los hombres con “junior college” o título de “Associates” en un programa académico, se denota por “AA-Académico” en el gráfico), son grafica-

<sup>7</sup> Históricamente hubo algunas diferencias entre estados en los sistemas educativos de los EE.UU: por ejemplo, Carolina del Sur sólo tenía tres años de escuela secundaria a principios del siglo XX.

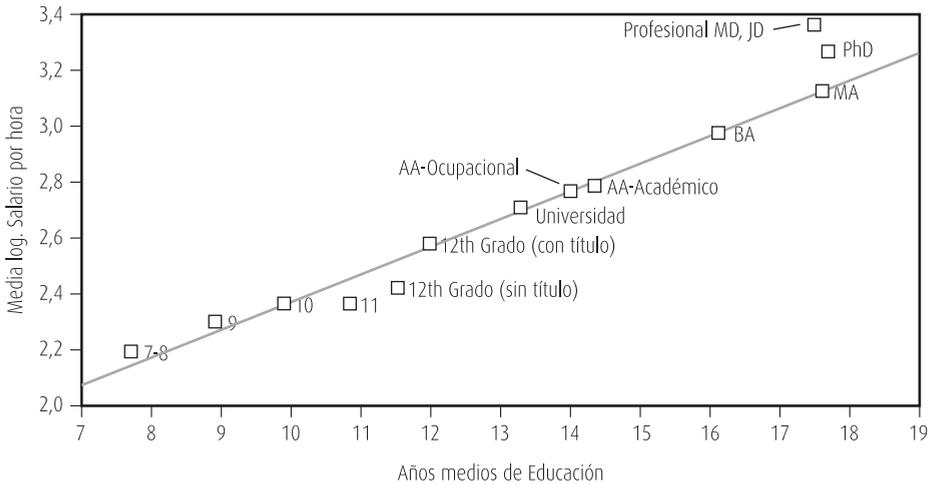
<sup>8</sup> Véase también Goodman (1979).

<sup>9</sup> Véase Park (1994, 1996) para un análisis adicional de la hipótesis de linealidad.

<sup>10</sup> Se toman los hombres en este rango de edad para abstraer el efecto de la experiencia. Como se observa en el gráfico 1.1a, después de los 40 años, los perfiles de edad-ingresos para diferentes grupos educativos, son aproximadamente paralelos.

Gráfico 1.2

RELACIÓN ENTRE EL LOGARITMO DEL SALARIO PROMEDIO POR HORA Y LA EDUCACIÓN TERMINADA, HOMBRES DE 40-45 AÑOS EN 1994-1996. EDUCACIÓN MEDIA POR CATEGORÍA DE GRADO ESTIMADO PARA LA CPS DE FEBRERO DE 1990



dos contra los años medios de educación para el grupo medido en la concordancia de febrero de 1990. Aparte de los hombres que reportan 11 años de educación, o 12 años sin título de secundaria, los datos para individuos entre 7 y 18 años de educación, se encuentran muy cerca de la línea que une a los graduados de secundaria y los graduados universitarios (superpuesta en el gráfico). Los dos grupos con mayor educación están también fuera de la línea. Mi hipótesis es que esto refleja censura en la variable años de educación, que solo fue reportada con un máximo de 18 años<sup>11</sup>. Basándose en los patrones del gráfico 1.2, puede ser razonable asignar una estimación de años terminados de educación a cada nivel educativo reportado y suponer una forma funcional lineal.

### ■ 1.2.3. ¿Qué medida de ingresos?

La literatura sobre la función de ingresos de capital humano ha analizado diversas medidas de ingreso —anual, semanal, por hora— casi siempre en forma logarítmica. La popularidad de la transformación logarítmica refleja varios aspectos. Primero, la distribución del loga-

<sup>11</sup> Las personas con un título en medicina o de derecho, por ejemplo, tienen al menos 20 años de educación, y muchos tienen más.

ritmo de los ingresos (especialmente el logaritmo del salario por hora) sorprendentemente se acerca mucho a una distribución normal. En igualdad de condiciones, muchos analistas preferirían, por tanto, el modelo de logaritmo de los ingresos. Otra razón práctica para emplear la transformación logarítmica es el aparente éxito de la función de ingresos de capital humano estándar (semilogarítmica). Como se demostró en la gráfica 1a,b, la distribución del logaritmo de los salarios entre la edad y los niveles educativos, se puede aproximar razonablemente bien con la suma de un término lineal de la educación y otro polinomial de la experiencia. Condicionado a la forma funcional del lado derecho de la ecuación (1), Heckman y Polachek (1974) investigaron transformaciones alternativas del ingreso, y concluyeron que la transformación logarítmica es la mejor del tipo Box-Cox. Finalmente, y quizá tan importante como cualquier otra consideración, la transformación logarítmica facilita la interpretación.

La elección del marco temporal sobre el cual medir los ingresos está frecuentemente condicionado por la necesidad: algunas bases de datos proporcionan información sobre los ingresos anuales, mientras otras lo hacen sobre el salario por hora o semanal. Dado que los individuos con mayor educación tienden a trabajar más, los rendimientos de la educación medidos serán mayores para ingresos semanales o anuales que para el salario por hora. Este hecho se ilustra en la tabla 1.1, que detalla los coeficientes de educación estimados en modelos análogos a la ecuación (1'), con los datos de ingresos y horas para hombres y mujeres en el CPS de marzo de 1994-1996. El cuestionario del CPS indaga sobre los ingresos del último año, el número total de semanas trabajadas en el último año y las horas habitualmente trabajadas por semana el último año. Por definición:

$$\text{Ingresos anuales} = \text{Ingresos por hora} \cdot \text{Horas semanales} \cdot \text{Semanas}$$

Cuando el logaritmo de los ingresos anuales es estimado como función de la educación y otras variables de control, el coeficiente de educación estimado es, por tanto, la suma de los coeficientes de educación para modelos paralelos que ajusten el logaritmo del ingreso por hora, el logaritmo de las horas por semana y el número de semanas por año. En el mercado laboral de EE.UU. a mediados de los noventa, aproximadamente dos tercios del rendimiento de la educación medido observado en datos de ingresos anuales, son atribuibles al efecto de la educación sobre el ingreso por hora, mientras que el resto es atribuible a los efectos sobre las horas por semana y las semanas por año (tabla 1.1).

#### ■ 1.2.4. Resumen

Esta breve visión general, sugiere que la función de ingresos de capital humano está viva y en buena forma. Un sencillo modelo de regresión con la especificación lineal de la educación y un polinomio de orden inferior en la experiencia potencial, explica entre el 20-35% de la variación en los datos de ingresos observados, con coeficientes con el signo y magnitud esperados y estimados con precisión en casi todas las aplicaciones. Una evaluación más detallada permite ver que el modelo es demasiado parsimonioso para caracterizar com-

Tabla 1.1

**COEFICIENTES DE EDUCACIÓN ESTIMADOS DE LA FUNCIÓN DE INGRESOS DEL CAPITAL HUMANO ESTÁNDAR AJUSTADA A LOS SALARIOS POR HORA, LOS INGRESOS ANUALES, Y VARIAS MEDIDAS DE HORAS PARA HOMBRES Y MUJERES EN MARZO 1994-1996, CPS\***

	VARIABLE DEPENDIENTE				
	LOG INGRESOS POR HORA (1)	LOG HORAS POR SEMANA (2)	LOG SEMANAS POR AÑO (3)	LOG HORAS AL AÑO (4)	LOG INGRESOS ANUALES (5)
<b>A. Hombres</b>					
Educación	0,100	0,018	0,025	0,042	0,142
coeficiente	(0,001)	(0,001)	(0,001)	(0,001)	(0,001)
R-cuadrado	0,328	0,182	0,136	0,222	0,403
<b>B. Mujeres</b>					
Educación	0,109	0,022	0,034	0,056	0,165
coeficiente	(0,001)	(0,001)	(0,001)	(0,001)	(0,001)
R-cuadrado	0,247	0,071	0,074	0,105	0,247

*Nota:* El cuadro presenta los coeficientes estimados del término lineal de la educación en el modelo lineal, que también incluye uno cúbico en la experiencia potencial y un indicador de la raza no blanca. Las muestras incluyen hombres y mujeres de 16 a 66 años que reportan ingresos salariales positivos en el año anterior. El salario por hora se construye dividiendo los ingresos salariales por el producto de las semanas trabajadas y las horas habituales por semana. Los datos de las personas cuyo salario sea inferior a \$ 2.00 o superior a \$ 150.00 (en dólares de 1995) son eliminados. Los tamaños muestrales son: 102.639 hombres y 95.309 mujeres.

pletamente la distribución conjunta de los ingresos, la edad y educación. No obstante, proporciona un punto natural de partida para construir modelos más complejos sobre la determinación de los salarios, y para investigar el efecto de otras variables como la raza, el género y las características de las empresas. Adicionalmente, el modelo convencional sirve como una referencia útil para la teorización acerca del efecto de la educación en el mercado de trabajo. Desde este punto de vista, la aproximación lineal de los ingresos con respecto a la educación, y la separabilidad del efecto de la educación y la experiencia son simplificaciones útiles que pueden facilitar la formulación de modelos teóricos manejables.

### ■ 1.3. MODELIZACIÓN CAUSAL DEL RENDIMIENTO DE LA EDUCACIÓN

#### ■ 1.3.1. Aspectos teóricos

La mayoría de aspectos teóricos subyacentes a la interpretación de los estudios recientes sobre la rentabilidad de la educación pueden ser ilustrados en el marco de un sencillo modelo estático que se basa en Becker (1967). De acuerdo con este modelo, cada individuo se enfrenta a un conjunto de oportunidades de mercado que da el nivel de ingreso aso-

ciado a elecciones alternativas de educación. Un modelo estático se abstrae de la naturaleza dinámica de los procesos de ingresos y educación y en su lugar se centra en la relación entre educación alcanzada e ingresos promedio a lo largo del ciclo vital. Este enfoque está justificado si las personas terminan su educación formal antes de entrar al mercado de trabajo (de forma no temporal o a tiempo parcial), y si el efecto de la educación sobre el logaritmo de los ingresos es separable del efecto de la experiencia, como se supone en la función de ingresos de capital humano estándar. De hecho la transición de la escuela al trabajo, es a menudo no lineal, dado que los adultos jóvenes van y vienen entre estudios a tiempo completo y parcial, y trabajo a tiempo parcial y completo<sup>12</sup>. No obstante, la mayoría de las personas han terminado su educación formal alrededor de los 25 años<sup>13</sup>.

Una versión analítica manejable del modelo de Becker se desarrolla en Card (1995a). Siguiendo esa presentación,  $y(S)$  representa el nivel de ingresos promedio (por año) que un individuo recibirá si adquiere el nivel de educación  $S$ .<sup>14</sup> Supongamos que un individuo escoge  $S$  para maximizar una función de utilidad  $U(S, y)$ , donde:

$$U(S, y) = \log y - h(s) \quad (2)$$

y  $h$  es una función convexa creciente. Esta función generaliza la función objetivo “valor presente actualizado” (VPA)

$$\int_0^{\infty} y(S) \exp(rt) dt = y(S) \exp(-rS)/r$$

que es apropiada si los individuos descuentan ingresos futuros a la tasa  $r$ , la educación es medida en años, y si suponemos que los individuos no obtienen ningún ingreso mientras están en la escuela, y obtienen  $y(S)$  por año a partir de entonces. La función objetivo VPA establece  $h(S) = rS$ . De manera más general, no obstante,  $h(S)$  puede ser estrictamente convexa si el coste marginal de cada año adicional de educación aumenta más que los ingresos no percibidos durante ese año, ya sea por consideraciones del mercado de crédito (Becker, 1967) o por preferencias<sup>15</sup>.

<sup>12</sup> Angrist y Newey (1991) estudian los cambios en los ingresos, asociados con incrementos en la educación adquirida después de que los hombres jóvenes entran al mercado de trabajo a tiempo completo.

<sup>13</sup> A la edad de 24 años, menos de una quinta parte de los adultos de los EE.UU. estaban matriculados en la escuela (incluso a tiempo parcial) a principio de los años noventa. Una sencilla tabulación de las tasas de matrícula por edad sugiere que la transición entre la escuela y el trabajo se ha vuelto más rápida en las últimas dos décadas, por lo menos, en los EE.UU. Por ejemplo, a pesar de que las tasas de matrícula de los de 20 años son mayores ahora que a finales de los setenta (47% de matriculados en 1992 frente a 37% en 1977) las tasas de matriculación de las personas antes de los treinta es menor en la actualidad (por ejemplo, un 7% para los de 30 años en 1992 frente a un 10% en 1977). Estas cifras son del CPS de octubre, e incluyen hombres y mujeres.

<sup>14</sup> El conjunto de oportunidades de mercado  $y(S)$  puede reflejar los efectos de la educación superior sobre la productividad, y/o otras fuerzas, tales como la señalización.

<sup>15</sup> Nótese que la Tasa Marginal de Sustitución (TMS) entre el ingreso y la educación es  $y(S)h'(S)$ . Bajo un criterio de valor presente actualizado, la TMS =  $r y(S)$ , ya que los costes de oportunidad del año  $S$ -ésimo de educación son los ingresos no percibidos  $y(S)$ . Si  $h'(S)$  es creciente en  $S$ , la TMS aumenta más rápido que  $y(S)$ .

Una elección óptima de educación, satisface la siguiente condición de primer orden:

$$h'(S) = y'(S)/y(S),$$

como se muestra en el gráfico 1.3. Una característica importante del tipo de funciones de preferencia definidas en la ecuación (2) es la linealidad en el logaritmo de los ingresos. Esto significa que las curvas de indiferencia en el gráfico 1.3 son paralelas verticalmente, con la consecuencia inmediata de que cualquier factor que aumente el logaritmo de los salarios para *todos* los niveles de educación, no tiene efectos sobre la elección de educación óptima. En principio esto no tiene por qué ser cierto. Por ejemplo, Griliches (1977) presenta una variante de preferencias de valor presente actualizada con la característica que un cambio uniforme al alza en el logaritmo de los ingresos para todos los niveles de educación conduce a la elección de un menor nivel educativo.

La heterogeneidad individual en la elección óptima de educación ilustrada en el gráfico 1.3 surge de dos fuentes: diferencias en los costes de (o la preferencia por) la educación, representadas por la heterogeneidad en  $h(S)$ ; y diferencias en los beneficios económicos de la educación, representados por la heterogeneidad en el rendimiento marginal de la educación  $y'(S)/y(S)$ . Una sencilla especificación de estos componentes de heterogeneidad es:

$$y'(S)/y(S) = b_i - k_1 S \quad (3a)$$

$$h'(S) = r_i - k_2 S \quad (3b)$$

Gráfico 1.3

### DETERMINACIÓN DEL NIVEL EDUCATIVO ÓPTIMO

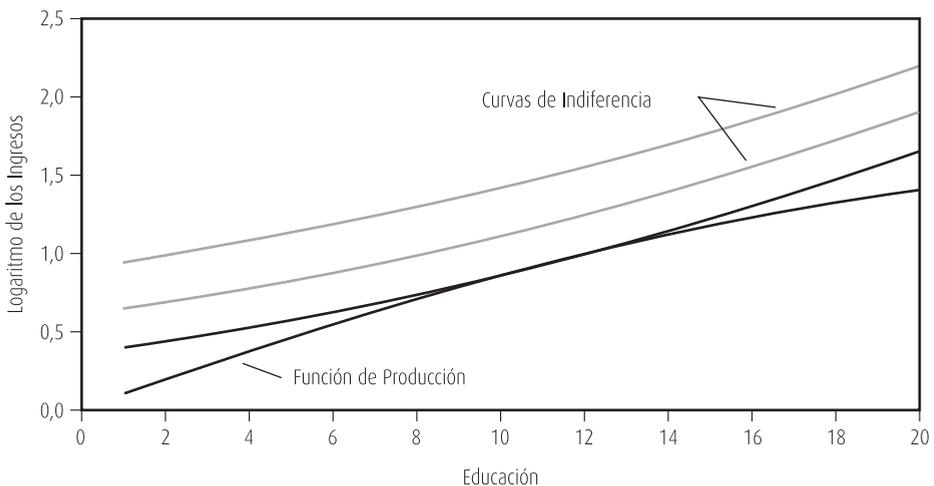
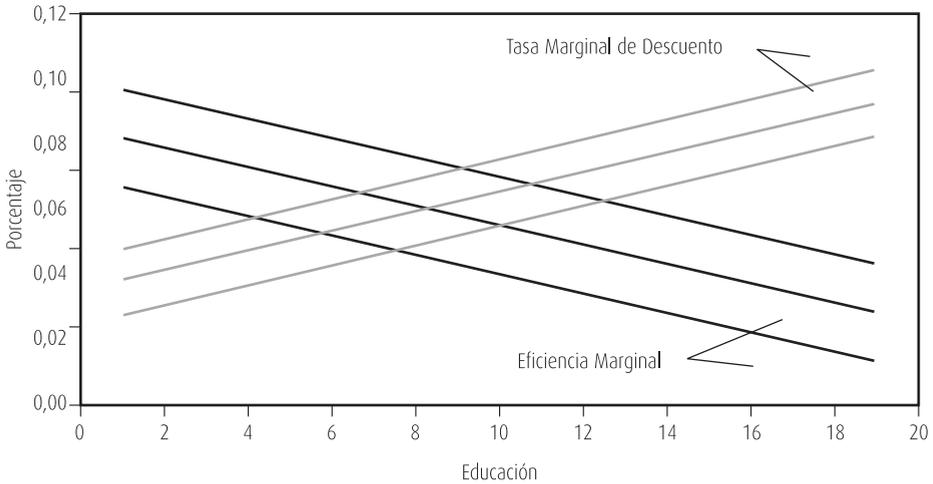


Gráfico 1.4

### REGISTRO DE COSTES MARGINALES Y BENEFICIOS MARGINALES PARA DIFERENTES INDIVIDUOS



donde  $b_i$  y  $r_i$  son variables aleatorias con medias  $\bar{b}$  y  $\bar{r}$ , y distribución conjunta en la población  $i = 1, 2, \dots$ , y  $k_1$  y  $k_2$  son constantes no negativas. Esta especificación implica que la elección óptima de educación es *lineal* en los términos de heterogeneidad específicos del individuo,

$$S_i^* = (b_i - r_i)/k \quad (4)$$

donde  $k = k_1 + k_2$ . El gráfico 1.4, ilustra la determinación de la educación óptima empleando las funciones del beneficio marginal y los costes marginales, descritos por las ecuaciones (3a) y (3b).

Dado que la educación formal es generalmente adquirida al principio de la vida, los individuos no necesariamente conocen los parámetros de sus funciones de ingreso, cuando toman su elección de educación. Por lo tanto,  $b_i$  debería ser interpretado como la mejor estimación que realiza el individuo, de los ingresos obtenidos por año de educación, a partir de sus primeros años de adulto. Se podría esperar que esta estimación variara menos entre individuos, que sus valores observados de educación. Además, la distribución de  $b_i$  puede cambiar en el tiempo, con cambios en las condiciones del mercado laboral, la tecnología, etcétera<sup>16</sup>. Por sim-

<sup>16</sup> Si cambios en el tiempo hacen que la rentabilidad media  $\bar{b}$  para una cohorte aumente o disminuya, pero dejan la distribución de  $b_i$ , en cualquier caso, inalterada, entonces los resultados que se presentan a continuación no se ven afectados.

plicidad, no obstante, trataré  $b_i$  como conocido al principio del ciclo vital y fijo en el tiempo: este supuesto probablemente lleva a exagerar el papel de la heterogeneidad de  $b_i$  en la determinación de la educación y el ingreso obtenido.

Al nivel de educación óptimo descrito en la ecuación (4), la tasa de rendimiento marginal de la educación para el individuo  $i$  es:

$$\beta_i = b_i - k_1 S_i^* = b_i (1 - k_1/k) + r_1 k_1/k$$

Incluso en este modelo simple, el equilibrio implica una *distribución* de rendimientos marginales en la población en la que al menos una de dos condiciones sea satisfecha: (a)  $r_i = \bar{r}$  para todo  $i$ , y  $k_2 = 0$  (es decir, curvas de indiferencia lineales con una pendiente uniforme  $\bar{r}$  en el gráfico 1.3); o (b)  $b_i = \bar{b}$  para todo  $i$  y  $k_1 = 0$  (es decir, conjuntos de oportunidad lineales con una pendiente uniforme  $\bar{b}$  en el gráfico 1.3).

En equilibrio general la distribución de los rendimientos marginales de la educación es endógena: una mayor oferta de trabajadores altamente educados presumiblemente disminuirá  $\bar{b}$ , y podría también afectar otras características de la distribución de  $b_i$ .<sup>17</sup> Desde el punto de vista de una cohorte de adultos jóvenes decidiendo sobre su educación, no obstante, la distribución del rendimiento de la educación se puede, razonablemente, considerar exógena. Por lo tanto, prefiero interpretar la ecuación (4) como una descripción de equilibrio parcial de la elección educativa relativa de una cohorte de adultos jóvenes, dado su *background* familiar, y el entorno institucional y las condiciones económicas que prevalecieron durante la adolescencia hasta los 25 años. Diferencias en los factores de *background* entre cohortes, dará lugar a mayor variación en la distribución de la rentabilidad marginal de la educación en la población en su conjunto.

### ■ 1.3.2. Educación observada e ingresos obtenidos

Para entender las implicaciones del anterior modelo para la educación observada y los ingresos obtenidos, obsérvese que la ecuación (3a) implica un modelo para el logaritmo de los ingresos de la forma:

$$\log y_i = \alpha_i + b_i S_i - \frac{1}{2} k_1 S_i^2$$

donde  $\alpha_i$  es una constante de integración específica del individuo. Esta es una versión algo más general de la forma funcional semilogarítmica adoptada por Mincer (1974), y cientos de estudios posteriores. En particular, la heterogeneidad individual afecta potencialmente, tanto al intercepto de la ecuación de ingresos (mediante  $\alpha_i$ ), como a la pendiente de la relación de educación (mediante  $b_i$ ). Por eso es conveniente reescribir esta ecuación como:

<sup>17</sup> Véase Freeman (1986) y Willis (1986) para una discusión de las implicaciones de equilibrio general en los modelos de educación óptima.

$$\log y_i = a_0 + \bar{b} S_i - \frac{1}{2} k_i S_i^2 + a_1 + (b_i - \bar{b}) S_i \quad (5)$$

donde  $\alpha_i \equiv \alpha_i - a_0$  tiene media 0. Las ecuaciones (4) y (5) describen un sistema de dos ecuaciones para la educación y los ingresos en términos de las variables aleatorias subyacentes  $a_i$ ,  $b_i$  y  $r_i$ .

A continuación, consideremos los ajustes lineales de  $a_i$  y  $(b_i - \bar{b})$  sobre la educación observada:

$$a_i = \lambda_0 (S_i - \bar{S}) + u_i, \quad (6a)$$

$$b_i - \bar{b} = \psi_0 (S_i - \bar{S}) + v_i, \quad (6b)$$

donde  $\bar{S}$  representa la educación media y  $E[S_i u_i] = E[S_i v_i] = 0$ . Los parámetros  $\lambda_0$  y  $\psi_0$  en las ecuaciones (6a) y (6b) son coeficientes de regresión teóricos:

$$\lambda_0 = \frac{\text{COV}(a_i, S_i)}{\text{var}(S_i)} = k \frac{\sigma_{ba} - \sigma_{ra}}{\sigma_b^2 - \sigma_r^2 - 2\sigma_{br}} \text{ y}$$

$$\psi_0 = \frac{\text{COV}(b_i, S_i)}{\text{var}(S_i)} = k \frac{\sigma_b^2 - \sigma_{br}}{\sigma_b^2 + \sigma_r^2 - 2\sigma_{br}}$$

donde  $\sigma_b^2$ ,  $\sigma_r^2$  y  $\sigma_{br}$  representan las varianzas y covarianza de  $b_i$  y  $r_i$ , y  $\sigma_{ba}$  y  $\sigma_{ra}$  representan las covarianzas de  $b_i$  y  $r_i$  respecto a  $a_i$ . Por simplicidad, supongamos que  $b_i$  y  $r_i$  tienen distribución conjunta simétrica<sup>18</sup>. Entonces, empleando la ecuación (A.3) y dado que la proyección lineal de  $S_i^2$  sobre  $S_i$  tiene pendiente  $2\bar{S}$ , se demuestra fácilmente que el límite de probabilidad del coeficiente de la regresión por Mínimos Cuadrados Ordinarios (MCO)  $b_{mco}$ , de la regresión del logaritmo de los ingresos sobre la educación, es:

$$\text{plim } b_{mco} = \bar{b} + \lambda_0 - k_1 \bar{S} + \psi_0 \bar{S} = \bar{\beta} + \lambda_0 + \psi_0 \bar{S} \quad (7)$$

donde  $\bar{\beta} \equiv E[\beta_i] = E[b_i - k_1 S_i] = \bar{b} - k_1 \bar{S}$ , es el rendimiento marginal promedio de la educación en la población<sup>19</sup>.

La ecuación (7) generaliza el análisis convencional del sesgo de habilidad en la relación entre educación e ingresos (véase Griliches, 1977)<sup>20</sup>. Supongamos que no existe heterogeneidad en los beneficios marginales de la educación (es decir,  $b_i = \bar{b}$ ) y que el logaritmo de los ingresos es lineal en la educación (es decir,  $k_1 = 0$ ). En este caso (7) implica que:

$$\text{plim } b_{mco} - \bar{b} = \lambda_0$$

<sup>18</sup> Este supuesto sugiere que  $E[(b_i - \bar{b})^3] = E[(r_i - \bar{r})^3] = E[(b_i - \bar{b})^3] = E[(r_i - \bar{r})(b_i - \bar{b})^2] = \dots = 0$ .

<sup>19</sup> Si las variables aleatorias  $r$  y  $b$  no están simétricamente distribuidas entonces la ecuación (7) contiene un término adicional igual a  $E[(b_i - \bar{b})(S_i - \bar{S})^2]$ . Véase el Apéndice A.

<sup>20</sup> A lo largo de este trabajo se utiliza el término "sesgo" para referirse a la diferencia entre el límite de probabilidad de un estimador y algún parámetro objetivo: generalmente, el rendimiento marginal medio educación de la educación en la población objeto de estudio.

ésta es la expresión estándar para el sesgo asintótico en el rendimiento de la educación estimado que surge de aplicar la fórmula de “variables omitidas” a un modelo de salarios con un coeficiente de educación constante  $\bar{b}$ . Según el modelo presentado aquí, el sesgo aparece debido a la correlación entre habilidad no observada  $a_i$  y el coste marginal de la educación  $r_i^{21}$ . Si los costes marginales son menores para los hijos de familias con *background* familiar más privilegiado, y si estos hijos tendieran además a ganar más a cualquier nivel de educación, entonces  $\sigma_{ra} < 0$ , lo que implica que  $\lambda_0 > 0$ .

Si tanto el intercepto como el coeficiente de la función de ingresos varía entre individuos, entonces la situación es más complicada. Como las personas con un mayor rendimiento de la educación tenderán a adquirir más educación, una regresión de corte transversal de los ingresos sobre la educación da lugar a una estimación sesgada al alza del rendimiento marginal promedio de la educación, incluso ignorando la variación en el intercepto de la función de ingresos. La magnitud de esta endogeneidad o sesgo de selección  $\psi_0 \bar{S}$ , depende de la importancia de la variación en  $b_i$  en la determinación de la varianza total de la educación.

Para verlo, se ha de tener en cuenta que la varianza de la educación es  $(\sigma_b^2 + \sigma_r^2 - 2\sigma_{br})/k^2$ . La parte de la varianza de la educación atribuible a diferencias en la pendiente de la relación ingresos-educación (en contraposición a las diferencias en los gustos o por el acceso a los recursos) se puede definir como:

$$f = \frac{\sigma_b^2 - \sigma_{br}}{\sigma_b^2 + \sigma_r^2 - 2\sigma_{br}}$$

Suponiendo que  $\sigma_{br} \leq 0$  (es decir, que los beneficios marginales de la educación no son mayores para personas con un mayor coste marginal de la educación), esta “proporción” está acotada entre cero y uno. El coeficiente de la regresión auxiliar definida en la ecuación (6b) es  $\psi_0 = kf \geq 0$ . Así, el componente de sesgo de endogeneidad en  $b_{mco}$  es:

$$\psi_0 \bar{S} = kf \bar{S} \geq 0$$

Incluso ignorando el tradicional término de sesgo de habilidad  $\lambda_0$ ,  $b_{mco}$  es, por tanto, un estimador sesgado al alza de  $\bar{\beta}$ ; sin embargo, cuanto mayor es  $f$ , mayor es el sesgo de endogeneidad.

Aparentemente, el modelo de ingresos especificado en la ecuación (5) parece inconsistente con la observación de que la relación de corte transversal entre el logaritmo del ingreso y la educación es aproximadamente lineal. Sin embargo, debido a la endogeneidad de la educación,  $S_i$  y  $(b_i - \bar{b})$  están correlacionados positivamente en la población, conduciendo a una relación convexa entre el logaritmo del ingreso y la educación en ausencia de cualquier concavidad en los conjuntos de oportunidad subyacentes. Más formalmente, la sustitución de (6a) y (6b) en la ecuación (5) lleva a:

<sup>21</sup> Como se señaló anteriormente, la forma de la ecuación (2) excluye una relación directa entre  $a_i$  y la elección de educación óptima.

$$\begin{aligned} \log y_i &= a_0 + \bar{b} S_i - \frac{1}{2} k_i S_i^2 + \lambda_0 (S_i - \bar{S}) + \psi_0 S_i (S_i - \bar{S}) + u_i + S_i v_i \\ &= c + (\bar{b} + \lambda_0 - \psi_0 \bar{S}) S_i + (\psi_0 - \frac{1}{2} k_i) S_i^2 + u_i + S_i v_i \end{aligned} \quad (5')$$

donde  $c$  es una constante. Si  $E[u_i | S_i] = E[v_i | S_i] = 0$  (supuestos algo más fuertes que las condiciones de ortogonalidad implícitas en las ecuaciones (6a) y (6b)), entonces la ecuación (5') implica que  $E[\log y_i | S_i]$  es una función cuadrática de la educación con un coeficiente de segundo orden  $(\psi_0 - 1/2k_i)$ . La relación empírica entre el logaritmo de los salarios y la educación será así aproximadamente lineal si y solo si  $k_i \approx 2\psi_0$ . Cuanto mayor sea la contribución de la variación de  $b_i$  sobre la varianza total de la educación, mayor es  $\psi_0$  y más convexa es la relación observada entre el logaritmo de los ingresos y la educación<sup>22</sup>.

### ■ 1.3.3. Errores de medida

Un problema importante en la literatura sobre rendimiento de la educación es el efecto de los errores de medición de la educación en las encuestas. Como hizo notar Griliches (1977, 1979) se esperaría que los errores de medida en la educación generen sesgos a la baja en cualquier estimador MCO de la relación entre educación e ingresos. Un supuesto convencional es que la educación *observada* ( $S_i^o$ ) difiere de la verdadera educación ( $S_i$ ) en un error aditivo:

$$S_i^o = S_i + \varepsilon_i,$$

con  $E[\varepsilon_i] = 0$ ,  $E[S_i \varepsilon_i] = 0$ , y  $E[\varepsilon_i^2] = \sigma_\varepsilon^2$ . Suponiendo que la ecuación (7) describe el límite de probabilidad de un estimador MCO utilizando la *verdadera* educación, el uso de la educación observada dará lugar a un estimador MCO con:

$$plim(b_{mco}) = R_o \{\bar{\beta} + \lambda_0 + \psi_0 \bar{S}\},$$

donde

$$R_o \equiv \text{cov}[S_i^o, S_i] / \text{var}[S_i^o] = \text{var}[S_i] / \{\text{var}[S_i] + \sigma_\varepsilon^2\}$$

es la fiabilidad de  $S_i^o$ , o la ratio señal-varianza total de la educación observada. Al considerar  $b_{mco}$  como un estimador de  $\bar{\beta}$  el sesgo asintótico es:

$$\text{Sesgo}_{mco} = R_o(\lambda_0 + \psi_0 \bar{S}) - (1 - R_o)\bar{\beta}$$

Las últimas tres décadas de investigaciones generalmente han encontrado que la fiabilidad de la educación autodeclarada es aproximadamente del 90%<sup>23</sup>, sugiriendo que el

<sup>22</sup> La observación de que la relación de corte transversal entre el logaritmo de los ingresos y la educación es aproximadamente lineal, no debería ser forzada al extremo. Dada la dispersión residual de los ingresos, una función cuadrática de la educación, con un término de segundo orden no trivial, también puede aparecer como lineal dentro del rango limitado de resultados educativos efectivamente observados en cualquier muestra.

<sup>23</sup> Véase, por ejemplo, Siegel y Hodge (1968), Miller *et al.* (1995), y Ashenfelter y Rouse (1998). Curiosamente, la limitada evidencia disponible sobre medidas administrativas de educación sugiere una ratio de fiabilidad similar, por ejemplo, Kane *et al.* (1997); Isacson (1997).

segundo término de esta expresión es del orden de  $-0,1\bar{\beta}$  en la mayoría de bases de datos. Dependiendo de las magnitudes de  $\lambda_0$  y  $\psi_0\bar{S}$  esto puede compensar parcialmente los sesgos presumiblemente positivos generados por las correlaciones entre la educación y los componentes de habilidad  $a_i$  y  $b_i$ .

El anterior argumento depende del supuesto de que los errores de medida de la educación no estén correlacionados con la verdadera educación. La educación generalmente se mide como una variable discreta con valores que oscilan entre límites superiores e inferiores fijos, no obstante, los errores en la educación declarada siguen probablemente un proceso de regresión a la media<sup>24</sup>. Específicamente, los individuos con niveles muy altos de educación no pueden declarar errores positivos en la educación, mientras que individuos con niveles muy bajos de educación no pueden declarar errores negativos en la educación. Si los errores de medida de la educación observada están correlacionados negativamente con la verdadera educación, la fiabilidad real de la medida de la educación observada puede ser ligeramente mayor que la fiabilidad estimada inferida a partir de la correlación entre dos medidas alternativas de educación<sup>25</sup>.

#### ■ 1.3.4. Estimaciones con variables instrumentales del rendimiento de la educación

Los científicos sociales ya hace tiempo que han reconocido que la correlación de corte transversal entre educación e ingresos puede diferir del verdadero efecto causal de la educación. Una solución estándar al problema de la inferencia causal es la utilización de variables instrumentales (VI): un investigador plantea la existencia de una variable observable que afecta la elección de educación pero que no está correlacionada con (o es independiente de) los factores de habilidad  $a_i$  y  $b_i$ . Por ejemplo, supongamos que el componente de coste marginal  $r_i$  está relacionado linealmente con un conjunto de variables  $Z_i$ :

$$r_i = Z_i\pi_1 + \eta_i$$

en este caso la ecuación de elección de educación sería:

$$S_i = Z_i\pi + (b_i - \eta_i)/k = \pi_0 + Z_i\pi + \xi_i \quad (4')$$

donde  $\pi = \pi_1/k$  y  $\xi_i = (b_i - \bar{\beta} - \eta_i)/k$ . En la literatura reciente se ha prestado mucha atención a lo que se podría llamar fuentes institucionales de variación de la educación, atribuible a características como la edad de escolarización mínima obligatoria, los costes de

<sup>24</sup> Este punto se plantea en un artículo reciente de Kane *et al.* (1997).

<sup>25</sup> Para ver esto, supongamos que hay dos medidas  $x_1$  y  $x_2$  de la verdadera cantidad  $x$ , con  $x_j = x + e_j$ , y supongamos que  $E[e_j|x] = -\alpha(x - \mu)$ , para  $j = 1, 2$ , donde  $\mu$  es la media de  $x$ . Descompongamos los errores de medición como,  $e_j = -\alpha(x - \mu) + v_j$ , y supongamos que los  $v_j$  son independientes entre sí y de  $x$ , y tienen varianzas iguales. La fiabilidad de  $x_1$  es  $R = \text{cov}[x, x_1]/\text{var}[x_1]$ . Tradicionalmente, la fiabilidad es medida por  $\rho = \text{cov}[x_1, x_2]/\text{var}[x_1]$  [(suponiendo que  $x_1$  y  $x_2$  tienen la misma varianza). Es fácil demostrar que  $\rho = (1 - \alpha)R$ .

matrícula de la educación superior o la proximidad geográfica a las escuelas. Estos factores institucionales ofrecen una oportunidad razonable de satisfacer los supuestos de exogeneidad estricta requeridos para que una variable se pueda considerar un buen instrumento.

En presencia de rendimientos educativos heterogéneos las condiciones requeridas para obtener un estimador por VI interpretable, son considerablemente más exigentes que las requeridas cuando la única fuente de sesgo de habilidad es la variación aleatoria en la constante de la ecuación de ingresos (es decir, la variación en  $a_i$ )<sup>26</sup>. Wooldridge (1997) presenta un útil análisis que puede ser directamente aplicado al sistema de las ecuaciones (4') y (5). Suponga por el momento que  $k_i = 0$  en la ecuación de ingresos, y consideremos tres supuestos adicionales sobre los componentes inobservables de (4') y (5):

$$E[\eta_i | Z_i] = 0, \quad E[a_i | Z_i] = 0, \quad E[(b_i - \bar{b}) | Z_i] = 0, \quad (9a)$$

$$E[(b_i - \bar{b})^2 | Z_i] = \sigma_{b_i}^2, \quad (9b)$$

$$E[\xi_i | b_i, Z_i] = \rho_i (b_i - \bar{b}). \quad (9c)$$

La ecuación (9a) supone que los componentes de heterogeneidad específicos del individuo son todos independientes de la media del instrumento  $Z$ . La ecuación (9b) dice que el segundo momento de  $b_i$  también es condicionalmente independiente de  $Z_i$ . Finalmente, la ecuación (9c) afirma que la esperanza condicional del componente no observado de la elección óptima de educación ( $\xi_i$ ) es lineal en  $b_i$ . Dado  $\xi_i \equiv (b_i - \bar{b} - \eta_i)/k$ , una condición suficiente para (9c), es que  $E[\eta_i | b_i, Z_i] = \rho_i (b_i - \bar{b})$ , en cuyo caso  $\rho_i = (1 - \rho)/k$ . Esto será cierto, por ejemplo, si  $b_i$  y  $\eta_i$  tienen una distribución normal bivalente que sea independiente de  $Z_i$ .

Bajo los supuestos (9a) - (9c), la esperanza condicional del componente residual de los ingresos atribuible a la heterogeneidad en  $b_i$  es:

$$\begin{aligned} E[(b_i - \bar{b}) S_i | Z_i] &= E[E[(b_i - \bar{b}) S_i | b_i, Z_i] | Z_i] = E[(b_i - \bar{b}) E[S_i | b_i, Z_i] | Z_i] \\ &= E[(b_i - \bar{b}) E[Z_i \pi + \xi_i | b_i, Z_i] | Z_i] = \rho_i \sigma_{b_i}^2 \end{aligned}$$

por lo tanto,

$$E[\log y_i | Z_i] = a_0 - \bar{b} Z_i \pi + \rho_i \sigma_{b_i}^2.$$

Así, el uso de  $Z_i$  como instrumento de la educación conducirá a estimaciones consistentes del rendimiento medio de la educación  $\bar{b}$  (pero estimaciones inconsistentes de  $a_0$ )<sup>27</sup>.

<sup>26</sup> Si el único componente individual específico de habilidad es  $a_i$ , entonces las ecuaciones (4') y (5) constituyen un sistema estándar de ecuaciones simultáneas y solo es necesario suponer que  $E[a_i | Z_i] = E[\eta_i | Z_i] = 0$ . La interpretación de las VI en presencia de coeficientes aleatorios es objeto de los trabajos de Angrist y Imbens (1995) y Angrist *et al.* (1996). Heckman y Vytlacil (1998) presentan algunos resultados similares a los aquí descritos.

<sup>27</sup> Los supuestos (9a) y (9b) no son los únicos que llevan a un estimador consistente de VI. Wooldridge propone como alternativa la par de supuestos:  $E[\xi_i^2 | Z_i] = \sigma_{\xi_i}^2$  y  $E[(b_i - \bar{b}) \xi_i | Z_i] = \tau \xi_i$ . La demostración de la consistencia del estimador VI entonces se deduce teniendo en cuenta que  $E[(b_i - \bar{b}) S_i | Z_i] = E[E[(b_i - \bar{b}) S_i | \xi_i, Z_i] | Z_i] = \tau \sigma_{\xi_i}^2$ .

Si los ingresos son una función cuadrática de la educación (es decir,  $k_1 > 0$ ), Wooldridge plantea que el cuadrado del valor predicho de la educación en la ecuación (4') se puede añadir a la lista de variables condicionantes y el argumento anterior sigue siendo válido.

Una alternativa estrechamente relacionada a la estimación por VI de un modelo de coeficientes aleatorios, es el enfoque de la función control, propuesta por primera vez en el contexto de la educación por Garen (1984). En lugar de las ecuaciones (9b) y (9c), suponemos que las esperanzas condicionales de  $a_i$  y  $b_i$  son lineales en  $S_i$  y  $Z_i$ :

$$E[a_i | S_i, Z_i] = \lambda_1 S_i + \lambda_2 Z_i, \quad (10a)$$

$$E[b_i - \bar{b} | S_i, Z_i] = \psi_1 S_i + \psi_2 Z_i, \quad (10b)$$

Como se explica en el Anexo A, manteniendo los supuestos  $E[a_i|Z_i] = E[b_i - \bar{b}|Z_i] = 0$ , estas condiciones son equivalentes a suponer:

$$E[a_i | S_i, Z_i] = \lambda_1 \xi_i, \quad (10a')$$

$$E[b_i - \bar{b} | S_i, Z_i] = \psi_1 \xi_i, \quad (10b')$$

donde  $\xi_i$  se define en la ecuación (4'). De ello se deduce inmediatamente que:

$$E[\log y_i | S_i, Z_i] = a_0 + \bar{b} S_i - 1/2 k_1 S_i^2 + \lambda_1 \xi_i + \psi_1 \xi_i S_i. \quad (11)$$

El enfoque de la función control para la estimación del rendimiento medio de la educación es sustituir el residuo estimado  $\xi_i$  en la forma reducida de la ecuación de educación (4') en lugar de  $\xi_i$  en la ecuación (11). Téngase en cuenta que la inclusión de  $\xi_i$  como un regresor adicional en la función de ingresos es numéricamente equivalente al enfoque de VI empleando  $Z_i$  como un instrumento de  $S_i$ . Bajo el supuesto de que  $E[a_i | Z_i] = 0$ , la inclusión de  $\xi_i$  en la función de ingresos estimada elimina el efecto de  $a_i$  sobre la relación observada entre el logaritmo de los ingresos y la educación. En general, no obstante, el enfoque VI estándar no eliminará la influencia de  $b_i$  sobre la covarianza entre la educación y los ingresos, a menos que  $E[(b_i - \bar{b})S_i | Z_i]$  sea independiente de  $Z_i$  (como en el caso bajo los supuestos de Wooldridge). Bajo el supuesto (10b) (o equivalentemente (10b')), la inclusión de  $\xi_i S_i$  como una segunda variable de control es suficiente para eliminar el sesgo de endogeneidad derivado de la correlación entre  $b_i$  y  $S_i$ . De este modo, el enfoque de la función de control se podría contemplar como una generalización de variables instrumentales.

### ■ 1.3.5. Limitaciones de las variables instrumentales

En ausencia de supuestos como los implícitos en las ecuaciones (9) o (10), incluso un estimador de variables instrumentales basado en un instrumento exógeno, no aportará necesariamente un estimador del rendimiento medio de la educación, insesgado asintóticamente. Para ilustrar este punto, consideremos la estimación VI utilizando el cambio en la educación asociado con una "reforma educativa" que lleva a una reducción proporcional del

coste marginal de la educación para estudiantes en un grupo específico de escuelas (o en una cohorte específica). Supongamos que la distribución conjunta de habilidades y preferencias  $(a_i, b_i, r_i)$  es la misma para individuos pertenecientes a las escuelas afectadas por la reforma (indexado por  $Z_i = 1$ ) y aquellos que no (indexado por  $Z_i = 0$ ), pero en las escuelas reformadas, la elección óptima de educación viene dada por:

$$S_i^* = (b_i - \theta r_i)/k, \quad (4'')$$

donde  $0 < \theta < 1$ . Evidentemente, las diferencias en  $Z_i$  están asociadas con diferencias en los niveles promedio de educación. Además, suponemos que las distribuciones de habilidad son iguales entre los estudiantes que atienden ambos grupos de escuelas. En este contexto, sin embargo, el efecto tratamiento de la reforma educativa es mayor para individuos que habrían tenido menores niveles de educación en ausencia de la reforma, provocando potenciales dificultades para la interpretación del estimador VI basado en  $Z_i$ :

Haciendo,  $r_i = \bar{r} + \eta_i$ , y observando que entre los individuos del grupo de comparación que asisten a escuelas no afectadas por la reforma,

$$S_i = (\bar{b} - \bar{r})/k + (b_i - \bar{b} - \eta_i)/k = \pi_0 + \xi_{0i},$$

mientras que entre los individuos del grupo de tratamiento que asistían a escuelas no reformadas:

$$S_i = (\bar{b} - \theta \bar{r})/k + (b_i - \bar{b} - \theta \eta_i)/k = \pi_1 + \xi_{1i},$$

Supongamos que  $E[\eta_i | b_i] = \rho(b_i - \bar{b})$ , entonces:

$$E[\xi_{0i} | b_i] = \rho_0 (b_i - \bar{b}),$$

donde  $\rho_0 = (1 - \rho)/k$ , mientras

$$E[\xi_{1i} | b_i] = \rho_1 (b_i - \bar{b}),$$

donde  $\rho_1 = (1 - \theta \rho)/k$ . Así, la correlación entre el error de la forma reducida de educación y la habilidad no observada es *diferente* en los grupos tratamiento y control, conduciendo a una violación de los supuestos requeridos por la estimación por VI, o por un estimador de la función control para ofrecer estimadores consistentes del rendimiento marginal medio de la educación.

La reforma educativa causa en un individuo dado (caracterizado por  $(a_i, b_i, \eta_i)$ ) un incremento en su educación en una cantidad

$$\Delta S_i = \pi_1 - \pi_0 + \eta_i (1 - \theta)/k.$$

El efecto (de primer orden) sobre los ingresos de los individuos es:

$$\Delta \log y_i = \beta_i \Delta S_i$$

Donde  $\beta_i$  es el rendimiento marginal  $i$ -ésimo a la educación en ausencia de intervención:

$$\beta_i = \bar{\beta} + b_i - \bar{b} - k_1(S_i - \bar{S}) = \bar{\beta} + (b_i - \bar{b})(1 - k_1/k) + \eta_i k_1/k,$$

Empleando estas expresiones, la esperanza del diferencial de ingresos entre los individuos del grupo tratamiento y control es:

$$E[\Delta \log y_i] = \bar{\beta}(\pi_1 - \pi_0) + k_1/k^2(1 - \theta) \sigma_\eta^2 + \sigma_{b\eta}(1 - \theta)(1 - k_1/k)/k.$$

donde las esperanzas son tomadas con respecto a la distribución conjunta de  $(a_i, b_i, \eta_i)$ . El estimador de VI de la rentabilidad de la educación basado en el instrumento  $Z_i$ ,  $b_{iv}$ , tiene límite de probabilidad:

$$\begin{aligned} \text{plim } b_{iv} &= \frac{E[\log y_{ip} | Z_i = 1] - E[\log y_i | Z_i = 0]}{E[S_i | Z_i = 1] - E[S_i | Z_i = 0]} \\ &= \bar{\beta} + \frac{1 - \theta}{k(\pi_1 - \pi_0)} \{ \sigma_\eta^2 k_1/k + \sigma_{b\eta}(1 - k_1/k) \}. \end{aligned}$$

Nótese que si es  $\eta_i$  una constante para todo  $i$  (en cuyo caso, todos los individuos obtienen el mismo incremento de educación), entonces  $\sigma_\eta^2 = \sigma_{b\eta} = 0$ , y el estimador VI es consistente para  $\bar{\beta}$ . De otro modo, suponiendo que  $\sigma_{b\eta} \leq 0$ , de manera que los individuos con mayores rendimientos educativos tienen mayor preferencia por la educación o menores tasas de descuento, el estimador VI puede estar positiva o negativamente sesgado respecto a  $\bar{\beta}$ . Un sesgo positivo surge porque el rendimiento marginal a la educación es decreciente en la educación si  $k_1 > 0$ : así personas con costes marginales de la educación inicialmente mayores tienden a tener mayores rendimientos marginales a un año adicional de educación. Lang (2003) llamó este fenómeno "sesgo de tasa de descuento". Por otro lado, surge un sesgo negativo porque los individuos con mayor coste marginal de educación, que son más afectados por la reforma educativa, tienen menores rendimientos marginales de la educación si  $\sigma_{b\eta} < 0$ . Es más probable que predomine un sesgo positivo, cuanto menor sea  $|\sigma_{b\eta}|$  respecto a  $\sigma_\eta^2$  y cuanto más cóncava sea la función de ingresos.

Para generalizar un poco este análisis, supongamos que la población puede ser dividida en grupos discretos de individuos ( $g = 1, 2, \dots$ ) que comparten valores comunes en los términos de habilidad latente y coste  $(a_g, b_g, \eta_g)$ . Consideremos una intervención (como un cambio en la edad de educación obligatoria mínima) que lleve a una variación  $\Delta S_g$  en la educación media del grupo  $g$ , y que  $\beta_g$  represente rendimiento marginal de la educación del grupo  $g$  en ausencia de intervención. Finalmente, supongamos que la intervención afecta a un grupo tratamiento de estudiantes que son idénticos a aquellos en el grupo comparación excepto por la intervención. En particular, suponemos que los individuos del grupo tratamiento y el grupo comparación con la misma habilidad latente y coste, tendrían la misma educación e ingresos en ausencia de intervención, y que las distribuciones conjuntas de habilidades y costes son las mismas en los dos grupos. Entonces un estimador VI del

rendimiento de la educación basado en un indicador del estatus de grupo tratamiento tendrá límite de probabilidad:

$$\text{plim} b_{IV} = \frac{E[\beta_g \Delta S_g]}{E[\Delta S_g]}$$

donde las esperanzas son tomadas con respecto a la distribución de probabilidad de la población entre celdas<sup>28</sup>. Nótese que si  $\Delta S_g \geq 0$  para todo  $g$  (que no tiene por qué ser cierto) entonces esta expresión puede ser interpretada como un promedio ponderado de los rendimientos marginales de la educación para cada grupo, con ponderación  $\Delta S_g$ <sup>29</sup>. Una condición necesaria y suficiente para  $\text{plim} b_{IV} = \bar{\beta}$  es  $E[\beta_g \Delta S_g] = E[\beta_g] E[\Delta S_g]$ . Entre las condiciones suficientes para esta igualdad están: (a)  $\beta_g = \bar{\beta}$  (rendimientos marginales idénticos para todos los grupos); o (b)  $E[\Delta S_g | \beta_g] = \Delta S$  (un efecto tratamiento aditivo homogéneo de la reforma educativa). En general, no obstante, si existe cierta heterogeneidad en la distribución de los rendimientos marginales de la educación, la estimación por VI basada en una intervención que afecta a un grupo reducido de población puede llevar a un rendimiento estimado por encima o por debajo del estimador MCO para la misma muestra.

Cabe destacar otros dos aspectos del estimador por variables instrumentales. Primero, el límite de probabilidad del estimador VI no está afectado por errores de medida de la educación<sup>30</sup>. Esto en sí mismo dará lugar a la tendencia del estimador de VI a superar el correspondiente estimador MCO del efecto de la educación sobre los ingresos. Segundo, la validez de un estimador VI depende crucialmente del supuesto de que los instrumentos están no correlacionados con otras características latentes de los individuos que puedan afectar sus ingresos. En el caso de un estimador VI basado en una variable indicador  $Z_i$ , por ejemplo, el estimador VI es numéricamente igual a la diferencia de la media del logaritmo de los ingresos entre el grupo  $Z_i = 1$  y el grupo  $Z_i = 0$ , dividido por la correspondiente diferencia en educación media<sup>31</sup>. Si la diferencia de educación es pequeña, *incluso pequeñas diferencias* de salarios medios entre los dos grupos son exacerbadas por el procedimiento de VI. Si  $Z_i$  fuera asignado aleatoriamente, como en un experimento real, esto no sería un problema. Sin embargo, en el caso de experimentos naturales o cuasi experimentos, las inferencias se basan en la diferencia entre grupos de individuos que asisten a la escuela en diferentes momentos, o en diferentes localizaciones, o diferencias en otras características tales

<sup>28</sup> Este análisis se puede generalizar permitiendo a las variables latentes tener diferentes distribuciones entre los grupos tratamiento y comparación. Esto puede conseguirse en principio "reponderando" el grupo comparación, aunque las ponderaciones pueden no ser directamente observables.

<sup>29</sup> Si  $\Delta S_g$  es dicotómica (de modo que el cambio en la educación es cero o tiene un efecto unitario), entonces el análisis anterior se puede situar en el contexto del "efecto tratamiento medio local" desarrollado por Angrist e Imbens (1995). Véase también Angrist *et al.* (1996).

<sup>30</sup> Esto supone que la variable instrumental no está correlacionada con el error de medida de la educación.

<sup>31</sup> Si se incluyen otras variables en el modelo, entonces las medias para cada submuestra se ajustan por el efecto de las variables.

como el mes de nacimiento. El uso de estas diferencias para hacer inferencias causales sobre el efecto de la educación requiere una muy cautelosa consideración de la hipótesis de que los grupos son idénticos más allá de las diferencias identificadoras anteriores.

### ■ 1.3.6. *Background* familiar

Aunque algunas de las investigaciones recientes más innovadoras sobre el valor de la educación han usado aspectos institucionales del sistema educativo para identificar el efecto causal de la educación, existe una larga tradición de usar la información del *background* familiar —como la educación del padre y de la madre— para controlar directamente la habilidad no observada o bien como una variable instrumental de la educación<sup>32</sup>. El interés por *background* familiar proviene del hecho de que la educación de los hijos está muy altamente correlacionada con las características de sus padres, y en particular con su nivel educativo<sup>33</sup>. La intensidad de esta correlación se muestra en la tabla 1.2, que presenta los coeficientes estimados de una sencilla regresión de los años terminados de educación sobre la educación del padre y la madre, utilizando muestras de cabezas de familia del General Social Survey de 1972-1996 (GSS)<sup>34</sup>. Para diferentes submuestras, cada año de educación adicional de cualquiera de los padres aumenta los años de educación terminados en cerca de 0,2 años, mientras que un aumento de un año en la educación promedio de los padres aumenta los años de educación terminados alrededor de 0,4 años. Aproximadamente el 30% de la variación observada en la educación entre los adultos de los EE.UU. se explica por la educación de los padres<sup>35</sup>.

A pesar de la elevada correlación intergeneracional con la educación, no está claro que las medidas de *background* familiar se puedan aceptar como instrumentos válidos de la educación, incluso aunque no tengan un efecto causal independiente sobre los ingresos. Para ilustrar este punto, supongamos por el momento que no existe heterogeneidad en el rendimiento de la educación (es decir,  $b_i = \bar{b}$ ) y dejemos de lado cualquier efecto de concavidad en la función del logaritmo de los ingresos (es decir, suponemos  $k_i = 0$ ). En este caso la ecuación (5) pasaría a tomar la forma:

$$\log y_i = a_0 + \bar{b}S_i + a_i, \quad (5'')$$

<sup>32</sup> Griliches (1979) presenta un panorama de estudios de educación y salarios utilizando *background* familiar.

<sup>33</sup> Véase Siebert (1985) para referencias de la literatura sobre *background* familiar y educación. Ashenfelter y Rouse (1998) muestran que hasta un 60% de la variación de corte transversal en los niveles educativos en su muestra de gemelos puede ser explicada por factores familiares (observables y no observables).

<sup>34</sup> Los modelos presentados en la tabla 2 incluyen controles de edad y año de nacimiento de los encuestados, aunque los coeficientes estimados (y los coeficientes R-cuadrado) no son muy diferentes, sin estos controles.

<sup>35</sup> Los resultados de la tabla 1.2 son bastante típicos de los encontrados en la literatura usando otras muestras. Si el *background* familiar es medido sólo por la educación de uno de los padres, el coeficiente se sitúa generalmente en un rango entre 0,3 y 0,4.

Tabla 1.2

## EFFECTO DE LA EDUCACIÓN DE LOS PADRES SOBRE LA EDUCACIÓN TERMINADA\*

		EDUCACIÓN DEL PADRE	EDUCACIÓN DE LA MADRE	R-CUADRADO
<b>Por raza y sexo</b>				
1.	Hombres blancos (N = 7330)	0,23 (0,01)	0,20 (0,01)	0,26
2.	Mujeres blancas (N = 8547)	0,20 (0,01)	0,21 (0,01)	0,32
3.	Hombres negros (N = 705)	0,18 (0,03)	0,22 (0,04)	0,33
4.	Mujeres negras (N = 1030)	0,09 (0,02)	0,22 (0,03)	0,28
<b>Hombres (todas las razas) por cohorte de nacimiento</b>				
5.	Nacidos antes 1920 (N = 430)	0,25 (0,05)	0,22 (0,05)	0,23
6.	Nacidos 1920-1934 (N = 1590)	0,26 (0,03)	0,24 (0,03)	0,22
7.	Nacidos 1935-1944 (N = 1785)	0,24 (0,02)	0,24 (0,02)	0,26
8.	Nacidos 1945-1954 (N = 2482)	0,22 (0,02)	0,19 (0,02)	0,23
9.	Nacidos 1955-1964 (N = 1593)	0,26 (0,02)	0,11 (0,02)	0,23
<b>Mujeres (todas las razas) por cohorte de nacimiento</b>				
10.	Nacidos antes 1920 (N = 492)	0,21 (0,04)	0,25 (0,04)	0,29
11.	Nacidos 1920-1934 (N = 1936)	0,19 (0,02)	0,25 (0,02)	0,28
12.	Nacidos 1935-1944 (N = 2112)	0,17 (0,02)	0,23 (0,02)	0,25
13.	Nacidos 1945-1954 (N = 2911)	0,19 (0,01)	0,18 (0,02)	0,25
14.	Nacidos 1955-1964 (N = 1960)	0,20 (0,01)	0,2 (0,02)	0,26

*Nota:* La variable dependiente en todos los modelos son los años de educación terminados. Las muestras incluyen personas entre 24 a 64 años en 1972-1996 de la *General Social Survey* con información válida sobre su educación y la de ambos padres. Los modelos en las filas 1-4 incluyen funciones cuadráticas de la edad y el año de nacimiento del encuestado, además de la educación del padre y de la madre. Los modelos de las filas 5-14 incluyen sólo un término lineal en la edad.

Consideremos una proyección lineal del componente de habilidad no observada sobre la educación individual y alguna medida de *background* familiar ( $F_i$ ):

$$a_i = \lambda_1 (S_i - \bar{S}) + \lambda_2 (F_i - \bar{F}) + u'_i, \quad (12)$$

Este ajuste bivalente se puede comparar al ajuste de  $a_i$  sobre  $S_i$  (es decir, la ecuación (6a) considerando otras dos regresiones auxiliares

$$F_i = \delta_0 + \delta_5 S_i + e_{1i}, \quad (13a)$$

$$S_i = \pi_0 + \pi_F F_i + e_{2i}, \quad (13b)$$

donde  $e_{1i}$  es ortogonal a  $S_i$  y  $e_{2i}$  es ortogonal a  $F_i$ . La fórmula convencional de variables omitidas implica que los coeficientes de las ecuaciones (6a) y (12) están relacionados por:

$$\lambda_0 = \lambda_1 + \lambda_2 \delta_5,$$

Además,  $\delta_5$  y  $\pi_F$  se relacionan con el coeficiente de correlación entre  $S_i$  y  $F_i$  ( $\rho_{SF}$ ) por

$$\delta_5 \pi_F = \rho_{SF}^2.$$

Empleando estos resultados es posible comparar tres estimadores potenciales de la ecuación (5''): el estimador MCO de una regresión univariante de los ingresos sobre la educación ( $b_{MCO}$ ); el estimador MCO de una regresión bivalente de los ingresos sobre la educación y el *background* familiar ( $b_{biv}$ ); y el estimador VI empleando  $F_i$  como instrumento de  $S_i$  ( $b_{iv}$ ). Los límites de probabilidad de estos tres estimadores son:

$$plim b_{MCO} = \bar{b} + \lambda_0 = \bar{b} + \lambda_1 + \lambda_2 \rho_{SF}^2 / \pi_F$$

$$plim b_{biv} = \bar{b} + \lambda_1$$

$$plim b_{iv} = \text{cov}[\log y_i, F_i] / \text{cov}[S_i, F_i] = \bar{b} + \lambda_1 + \lambda_2 / \pi_F$$

Además, el límite de probabilidad del coeficiente de  $F_i$  en la regresión bivalente es  $\lambda_2$ <sup>36</sup>. Suponiendo que  $\lambda_1 \geq 0$ ,  $\lambda_2 \geq 0$ , y  $\pi_F > 0$ ,

$$\bar{b} \leq plim b_{biv} \leq plim b_{MCO} \leq plim b_{iv}.$$

Si  $a_i$  y  $S_i$  no están correlacionadas al controlar por  $F_i$ , entonces  $\lambda_1 = 0$  y el estimador MCO bivalente es consistente para  $\bar{b}$ . Si no, los tres estimadores probablemente estarán sesgados al alza, con sesgo mayor en los estimadores univariantes por MCO y VI que en el estimador bivalente a menos que  $\lambda_2 = 0$ <sup>37</sup>.

<sup>36</sup> Si  $F_i$  tiene un efecto causal independiente  $\gamma$  sobre los ingresos, entonces la ecuación (5'') incluye un término  $\gamma F_i$ . En este caso el límite de probabilidad del coeficiente de regresión de  $F_i$  es  $\gamma + \lambda_2$ , y  $plim b_{iv}$  incluye un componente  $\gamma / \pi_F$ .

<sup>37</sup> Supongamos que la medida de *background* familiar es la educación promedio de la madre y del padre. Los resultados de la tabla 1.2 sugieren que  $\pi_F = 0,4$  y  $\rho_{SF}^2 = 0,3$ , lo que implica que el estimador de MCO univariante superará el MCO bivalente por cerca de un  $0,75 \lambda_2$ , mientras que el estimador de VI superará el estimador MCO bivalente en  $2,5 \lambda_2$ .

Este análisis se puede ampliar fácilmente al caso en el que  $b_i$  varía entre individuos. Supongamos que los ingresos están dados por la ecuación (5) y consideremos el ajuste de  $b_i$  sobre  $S_i$  y  $F_i$ :

$$b_i - \bar{b} = \psi_1 (S_i - \bar{S}) + \psi_2 (F_i - \bar{F}) + v_i', \quad (14)$$

Como en el caso de los coeficientes  $\lambda_0$  y  $\lambda_1$ , los coeficientes  $\psi_0$  en la ecuación (6b) y  $\psi_1$  en la ecuación (14) estar relacionados por:

$$\psi_0 = \psi_1 + \psi_2 \delta_S = \psi_1 + \psi_2 \rho_{SF}^2 / \pi_F$$

Utilizando la ecuación (A.3) del Anexo, y suponiendo que  $b_i$ ,  $S_i$  y  $F_i$  siguen una distribución conjunta simétrica, es fácil demostrar que:

$$\begin{aligned} \text{plim } b_{mco} &= \bar{\beta} + \lambda_0 + \psi_0 \bar{S} = \bar{\beta} + \lambda_1 + \psi_1 \bar{S} + (\lambda_2 + \psi_2 \bar{S}) \rho_{SF}^2 / \pi_F, \\ \text{plim } b_{biv} &= \bar{\beta} + \lambda_1 + \psi_1 \bar{S}, \\ \text{plim } b_{iv} &= \bar{\beta} + \lambda_1 + \psi_1 \bar{S} + (\lambda_2 + \psi_2 \bar{S}) / \pi_F. \end{aligned}$$

Además, el límite de probabilidad del coeficiente de  $F_i$  en la regresión bivalente es  $\lambda_2 + \psi_2 \bar{S}$ . En presencia de heterogeneidad en  $b_i$ , es posible reinterpretar  $\lambda_1$  como  $(\lambda_1 + \psi_1 \bar{S})$  y  $\lambda_2$  como  $(\lambda_2 + \psi_2 \bar{S})$ . Suponiendo que  $\lambda_1 + \psi_1 \bar{S} \geq 0$ ,  $\lambda_2 + \psi_2 \bar{S} \geq 0$  y  $\pi_F > 0$ , los límites de probabilidad de los tres estimadores continúan satisfaciendo las desigualdades,

$$\bar{\beta} \leq \text{plim } b_{biv} \leq \text{plim } b_{mco} \leq \text{plim } b_{iv}.$$

En resumen, a menos que  $\lambda_1 = \lambda_2 = \psi_1 = \psi_2 = 0$  en las ecuaciones de ajuste de de habilidad individual  $a_i$  y  $b_i$ , el *background* familiar no es un instrumento válido de la educación, incluso aunque el *background* familiar no tenga un efecto causal directo sobre los ingresos. La inclusión de controles de *background* familiar puede reducir los sesgos en el rendimiento medido, pero aún puede inducir una estimación sesgada al alza del rendimiento marginal medio a menos que *todos* los componentes de la habilidad no observada sean absorbidos por los controles de *background* familiar (es decir, a menos que  $\lambda_1 = \psi_1 = 0$ ). Finalmente, obsérvese que en el caso especial en que  $\lambda_1 + \psi_1 \bar{S} = \lambda_2 + \psi_2 \bar{S}$ , el sesgo al alza en el coeficiente de educación estimado del modelo bivalente que controla el *background* familiar es igual al límite de probabilidad del coeficiente de la variable de *background* familiar. En estas circunstancias, es posible recuperar un estimador insesgado del rendimiento marginal medio de la educación restando el coeficiente del *background* familiar del coeficiente de la educación. Esto es equivalente a un estimador "*intra-familia*" y será discutido con mayor detalle en la siguiente sección.

El análisis precedente supone que la verdadera educación es observable. En el caso más realista en el que sólo se dispone de una medida imprecisa del nivel educativo, una comparación entre los tres estimadores debe tener en cuenta el impacto diferencial de los errores de medida sobre el MCO univariante, el MCO bivalente, y el VI. Sea  $R_0$  la fiabilidad

de la medida de educación y supongamos, por el momento, que  $F_i$  se mide sin error. Como se vio antes, el estimador MCO univariante es atenuado por el factor  $R_0$ :

$$plim b_{mco} = R_0[\bar{\beta} + \lambda_1 + \psi_1 \bar{S} + (\lambda_2 + \psi_2 \bar{S}) \rho_{SF}^2 / \pi_F].$$

La adición de  $F_i$  al modelo de ingresos tenderá a conducir a una mayor atenuación del coeficiente de educación medido, ya que parte de la verdadera educación puede ser inferida de  $F_i$ . Como se mostró en el Anexo A, el estimador MCO bivalente es atenuado por el factor  $R_1$ :

$$plim b_{biv} = R_1[\bar{\beta} + \lambda_1 + \psi_1 \bar{S}].$$

donde  $R_1 = (R_0 - \rho_{SF}^2) / (1 - \rho_{SF}^2) < R_0$ . Por ejemplo, si  $R_0 \approx 0,9$  y  $\rho_{SF}^2 \approx 0,3$  entonces  $R_1 \approx 0,85$ . A diferencia de cualquiera de los estimadores MCO, el estimador por VI no se ve afectado por el error de medida. Por lo tanto, si  $F_i$  es medido sin error, los errores de medida de la educación tenderán a reforzar la ordenación esperada de los estimadores MCO univariante, MCO bivalente y por VI, al introducir el mayor sesgo de atenuación en el estimador MCO bivalente, un sesgo intermedio en el estimador MCO univariante, mientras que el estimador VI se mantiene insesgado.

En muchas bases de datos la información del *background* familiar es obtenida de niños o bien retrospectivamente de adultos. En cualquier caso, se podría esperar que contenga errores considerables de información. De hecho, Ashenfelter y Rouse (1998, anexo 1) hallan que la fiabilidad de la información ofrecida por gemelos sobre la educación de sus madres es aproximadamente del 80%, comparado con una ratio de fiabilidad del 90% para su propia educación. La presencia de errores de medida en  $F_i$  introduce una expresión más compleja para el límite de probabilidad del estimador MCO bivalente. Concretamente, la fórmula del error de medida presentado en el Apéndice implica que,

$$plim b_{biv} = R_1[\bar{\beta} + \lambda_1 + \psi_1 \bar{S}] + (\lambda_2 + \psi_2 \bar{S}) (1 - R_F) \rho_{SF}^2 / (\pi_F (1 - \rho_{SF}^2)),$$

donde  $R_F$  es la fiabilidad del *background* familiar medido. El segundo término en esta expresión es 0 si el verdadero coeficiente del *background* familiar en el modelo bivalente es 0 (es decir, si  $\lambda_2 + \psi_2 \bar{S} = 0$ ), o si  $R_F = 1$ . Sin embargo, si el verdadero coeficiente de  $F_i$  es positivo y  $\pi_F > 0$ , entonces los errores de medida de  $F_i$  inducen un sesgo positivo en el coeficiente de educación que podría compensar parcialmente el efecto directo de atenuación del error de medida de  $S_i$ . Por ejemplo, si  $R_F \approx 0,8$ ,  $\rho_{SF}^2 \approx 0,3$  y  $\pi_F \approx 0,4$ , el segundo término es del orden del 20% del verdadero coeficiente del *background* familiar.

### ■ 1.3.7. Modelos para hermanos y gemelos

Una alternativa al enfoque de variables instrumentales para el problema de inferencia causal es estudiar la educación y los ingresos para hermanos, gemelos, o pares padre-

hijo/madre-hija. La idea central tras esta estrategia es que parte de las diferencias no observadas que sesgan la comparación de corte transversal de la educación y los ingresos se reducen o eliminan dentro de las familias<sup>38</sup>. Por ejemplo, supongamos que dos observaciones (con subíndice  $j = 1$  o  $2$ ) están disponibles para cada familia (con subíndice  $i$ ), y que los ingresos del individuo  $j$  de la familia  $i$  son generados por:

$$\log y_{ij} = a_0 + \bar{b} S_{ij} - \frac{1}{2} k_i S_{ij} + a_{ij} + (b_{ij} - \bar{b}) S_{ij}, \quad (15)$$

Un modelo de “efectos familiares puros” implica que  $a_{ij} = a_i$  y  $b_{ij} = b_i$ . Consideremos los ajustes lineales de  $a_i$  y  $b_i - \bar{b}$  sobre la educación observada de los dos miembros de la familia:

$$a_i = \lambda_1 (S_{i1} - \bar{S}_1) + \lambda_2 (S_{i2} - \bar{S}_2) + u_i, \quad (16a)$$

$$b_i - \bar{b} = \psi_1 (S_{i1} - \bar{S}_1) + \psi_2 (S_{i2} - \bar{S}_2) + v_i, \quad (16b)$$

Suponiendo que  $b_i$ ,  $S_{i1}$  y  $S_{i2}$  siguen una distribución conjunta simétrica, la ecuación (A.3) en el Anexo A sugiere que los ingresos observados de los miembros de la familia están relacionados con su educación por:

$$\log y_{i1} = c_1 + (\bar{\beta} + \lambda_1 + \psi_1 \bar{S}_1) S_{i1} + (\lambda_2 + \psi_2 \bar{S}_1) S_{i2} + e_{i1}, \quad (17a)$$

$$\log y_{i2} = c_2 + (\lambda_1 + \psi_1 \bar{S}_2) S_{i1} + (\bar{\beta} + \lambda_2 + \psi_2 \bar{S}_1) S_{i2} + e_{i1}, \quad (17b)$$

donde  $c_1$  y  $c_2$  son constantes y los residuos  $e_{ij}$  son ortogonales respecto a  $S_{i1}$  y  $S_{i2}$ . Las ecuaciones (17a) y (17b) constituyen un sistema de regresiones aparentemente no relacionadas<sup>39</sup>. Dado que no existen restricciones de exclusión, el sistema puede ser estimado de manera eficiente mediante la aplicación de MCO a cada ecuación. Alternativamente, se puede construir la diferencia *intra-familia* en el logaritmo de los ingresos  $\Delta \log y_i = \log y_{i1} - \log y_{i2}$ , y considerar un modelo de la forma:

$$\Delta \log y_i = \mu_1 S_{i1} + \mu_2 S_{i2} + e_i. \quad (18)$$

Numéricamente, la estimación MCO de los coeficientes de (18) será igual a la *diferencia* en las correspondientes de estimaciones MCO de los coeficientes en (17a) y (17b)<sup>40</sup>.

El atractivo del modelo de “efectos familiares puros” proviene del hecho de que potencialmente se pueden recuperar las estimaciones de  $\bar{\beta}$  de las diferencias en los coefi-

<sup>38</sup> Por supuesto, un estimador *intra-familia* puede ser objeto de una interpretación como VI: el instrumento de la educación es la desviación de la educación de un individuo respecto de la media de su familia.

<sup>39</sup> Un sistema de ecuaciones como (17a) y (17b) es a veces llamado especificación de “efectos aleatorios correlacionados”. La idea de proyectar el componente no observable de los residuos (es decir, el efecto familia) sobre los valores observados de los gemelos, y luego sustituir la ecuación predicha de nuevo en la ecuación de ingresos, fue popularizada por Chamberlain (1982).

<sup>40</sup> Si otras variables  $X_{ij}$  son incluidas en el modelo, entonces el modelo en primeras diferencias debe contener  $X_{i1}$  y  $X_{i2}$ , para que la condición de aditividad se mantenga.

cientes de la ecuación (17), o de los coeficientes de la ecuación (18). Por ejemplo, supongamos que no existe heterogeneidad en  $b_i$ . En este caso  $\psi_1 = \psi_2 = 0$  en las ecuaciones (17a) y (17b), y, por tanto, el coeficiente en la ecuación (18) satisface:

$$\text{plim } \mu_1 = \text{plim } -\mu_2 = \bar{\beta}$$

Por lo tanto, un contraste de la hipótesis  $\mu_1 = -\mu_2$  ofrece un test de especificación del modelo del modelo de "efectos familiares puros" cuando la heterogeneidad en la pendiente de la educación  $b_i$  no se considera<sup>41</sup>.

Un modelo de "efectos familiares puros" es particularmente plausible para gemelos idénticos, dado que los gemelos idénticos comparten genética, y casi siempre comparten el mismo entorno de *background* familiar. Para gemelos idénticos, también parece natural imponer las condiciones de simetría  $\lambda_1 = \lambda_2 = \lambda$ ,  $\psi_1 = \psi_2 = \psi$ , y  $\bar{S}_1 = \bar{S}_2 = \bar{S}$ , dado que la identidad de unos gemelos específicos es arbitraria. Con estas simplificaciones las ecuaciones (17a) y (17b) se reducen a

$$\begin{aligned} \log y_{i1} &= c_1 + (\bar{\beta} + \lambda + \psi \bar{S}) S_{i1} + (\lambda + \psi \bar{S}) S_{i2} + e_{i1}, \\ &= c_1 + \bar{\beta} S_{i1} + (\lambda + \psi \bar{S}) (S_{i1} + S_{i2}) + e_{i1}, \end{aligned} \quad (17a')$$

$$\begin{aligned} \log y_{i2} &= c_2 + (\lambda + \psi \bar{S}) S_{i1} + (\bar{\beta} + \lambda + \psi \bar{S}) S_{i2} + e_{i2}, \\ &= c_2 + \bar{\beta} S_{i2} + (\lambda + \psi \bar{S}) (S_{i1} + S_{i2}) + e_{i2}, \end{aligned} \quad (17b')$$

Estas ecuaciones expresan el logaritmo de los ingresos de un gemelo en términos de su propia educación y de la educación total (o promedio) de la pareja<sup>42</sup>. Bajo los supuestos de una especificación de "efectos familiares puros", todos los sesgos derivados de las correlaciones entre la habilidad no observada y la educación son absorbidos por el coeficiente de la educación total o promedio de la familia, y el coeficiente de la educación propia ofrece un estimador insesgado del rendimiento marginal medio de la educación. (Esta estimación es equivalente numéricamente a restar el coeficiente de educación estimado del hermano del coeficiente de la propia educación). Téngase en cuenta que si los supuestos de efectos familiares puros y de simetría son satisfechos, se puede estimar  $\bar{\beta}$  con datos de ingresos de un solo uno de los gemelos, siempre que se conozca los niveles de educación de ambos<sup>43</sup>.

En el caso de hermanos o pares de padres-hijos, puede ser menos verosímil que individuos de la misma familia tengan exactamente los mismos parámetros de habilidad. Por ejemplo, los hermanos mayores pueden ser tratados de modo diferente que los más jóve-

<sup>41</sup> Estos tests han sido ampliamente empleados en otras aplicaciones del modelo de efectos aleatorios correlacionados: por ejemplo, Jakubson (1988).

<sup>42</sup> Ecuaciones similares son derivadas por Ashenfelter y Rouse (1998).

<sup>43</sup> Los argumentos de intercambiabilidad sugieren que se debe cumplir la simetría para una ordenación aleatoria de los gemelos de cada familia. Sin embargo, si la muestra de "gemelos 1" está condicionada al empleo y algunos de los individuos de la muestra de "gemelos 2" no trabajan, el orden ya no es aleatorio, y la simetría podría no ser una restricción válida.

nes, dando lugar a diferencias en los resultados potenciales de mercado de trabajo<sup>44</sup>. Los supuestos del modelo de “efectos familiares puros” pueden ser relajados de la siguiente manera. Consideremos los ajustes lineales,

$$a_{i1} = \lambda_{11} (S_{i1} - \bar{S}_1) + \lambda_{12} (S_{i2} - \bar{S}_2) + u_{i1}, \quad (19a)$$

$$a_{i2} = \lambda_{21} (S_{i1} - \bar{S}_1) + \lambda_{22} (S_{i2} - \bar{S}_2) + u_{i2}, \quad (19b)$$

$$b_{i1} - \bar{b} = \psi_{11} (S_{i1} - \bar{S}_1) + \psi_{12} (S_{i2} - \bar{S}_2) + v_{i1}, \quad (19c)$$

$$b_{i2} - \bar{b} = \psi_{21} (S_{i1} - \bar{S}_1) + \psi_{22} (S_{i2} - \bar{S}_2) + v_{i2}, \quad (19d)$$

donde  $u_{ij}$  y  $v_{ij}$  son ortogonales respecto a  $S_{i1}$  y  $S_{i2}$ . Para hermanos o mellizos ordenados aleatoriamente, es natural suponer que los coeficientes satisfagan las restricciones de simetría:  $\lambda_{11} = \lambda_{22}$ ,  $\lambda_{12} = \lambda_{21}$ ,  $\psi_{11} = \psi_{22}$ , y  $\psi_{12} = \psi_{21}$ , aunque para pares padres-hijos, madres-hijas, estos supuestos son menos atractivos<sup>45</sup>. Sustituyendo estas ecuaciones en el modelo de ingresos (15) y considerando la proyección lineal sobre las variables de educación observada, se obtiene una versión generalizada de las ecuaciones (17a) y (17b)<sup>46</sup>:

$$\log y_{i1} = c_1 + \tau_{11} S_{i1} + \tau_{12} S_{i2} + e_{i1}, \quad (20a)$$

$$\log y_{i2} = c_2 + \tau_{21} S_{i1} + \tau_{22} S_{i2} + e_{i2}, \quad (20b)$$

donde

$$\begin{aligned} \tau_{11} &= \bar{\beta} + \lambda_{11} + \psi_{11} \bar{S}_1 & \tau_{12} &= \lambda_{12} + \psi_{12} \bar{S}_1 \\ \tau_{21} &= \lambda_{21} + \psi_{21} \bar{S}_1 & \tau_{22} &= \bar{\beta} + \lambda_{22} + \psi_{22} \bar{S}_2 \end{aligned}$$

Evidentemente,  $\bar{\beta}$  no es identificable a partir de los coeficientes de la regresión aparentemente no relacionada en las ecuaciones (20a) y (20b), incluso con el supuesto de simetría *intra-familia*, aunque si  $\psi_{ij} = 0$  o  $\bar{S}_1 = \bar{S}_2$  entonces la simetría impone dos restricciones lineales sobre los coeficientes ( $\tau_{11} = \tau_{22}$  y  $\tau_{21} = \tau_{12}$ ).

No obstante, puede ser posible fijar un límite superior en el rendimiento marginal medio de la educación usando datos de mellizos o hermanos. Específicamente, supongamos que  $\lambda_{11} \geq \lambda_{12}$  y  $\psi_{11} \geq \psi_{12}$ ; aproximadamente, este supuesto significa que la educación del individuo 1 es más informativa de su habilidad, que la educación del individuo 2<sup>47</sup>. En este caso,

<sup>44</sup> Véase Kessler (1991). Kessler llega a la conclusión de que el orden de nacimiento tiene poco o ningún efecto en los resultados económicos una vez que se tiene en cuenta el tamaño familiar.

<sup>45</sup> Para los pares padre-hijo, Ashenfelter y Zimmerman (1997) proponen un modelo ligeramente generalizado en el que  $\lambda_{2j} = \alpha \lambda_{1j}$ . Ignoran la heterogeneidad en  $b_i$ .

<sup>46</sup> Nótese que sigo suponiendo que  $(b_{ij}, S_{i1}, S_{i2})$  siguen una distribución conjunta simétrica.

<sup>47</sup> Los supuestos sobre las magnitudes relativas de los coeficientes de las proyecciones, son más naturales si  $S_{i1}$  y  $S_{i2}$  tienen las mismas varianzas. En ese caso,  $\lambda_{11} - \lambda_{12} = A (\text{cov}[a_{i1}, S_{i1}] - \text{cov}[a_{i1}, S_{i2}])$  para algún coeficiente  $A$  positivo; una expresión similar aplica para  $\psi_{11} - \psi_{12}$ .

$$\begin{aligned} \text{plim } \tau_{11} - \tau_{12} &= \bar{\beta} + \lambda_{11} + \psi_{11} \bar{S}_1 - \lambda_{12} - \psi_{12} \bar{S}_1 \\ &= \bar{\beta} + (\lambda_{11} - \lambda_{12}) + (\psi_{11} - \psi_{12}) \bar{S}_1 \\ &\geq \bar{\beta}, \end{aligned}$$

por lo que un estimador del límite superior de  $\bar{\beta}$  es  $\tau_{11} - \tau_{12}$ , que es la diferencia entre el efecto de la propia educación y el efecto de la educación del otro miembro de la familia en una ecuación ingresos de uno de los miembros de la familia<sup>48</sup>. Mecánicamente, esta diferencia es igual al coeficiente de la propia educación cuando la educación promedio de la familia es incluida en la regresión, como en la ecuación (17a')<sup>49</sup>.

Desafortunadamente, no hay ninguna garantía de que esta cota sea más restrictiva que la cota que implica el estimador MCO de corte transversal. En otras palabras, es posible que el estimador MCO sufra de un menor sesgo al alza que el estimador *intra-familiar* basado en la ecuación (17a). Una condición necesaria y suficiente para que el estimador *intra-familia* tenga un menor sesgo asintótico es:

$$|\lambda_0 + \psi_0 \bar{S}_1| > |(\lambda_{11} - \lambda_{12} + (\psi_{11} - \psi_{12}) \bar{S}_1)|,$$

donde  $\lambda_0$  y  $\psi_0$  son los coeficientes definidos en la ecuación (6a) y (6b). Para ilustrar los problemas subyacentes a la comparación entre los estimadores MCO y *intra-familia*, supongamos no existencia de heterogeneidad en el intercepto  $a_{ij}$  de la función de ingresos, de modo que el sesgo asintótico relativo de los estimadores MCO y *intra-familia* dependen de la comparación entre  $\psi_0$  y  $\psi_{11} - \psi_{12}$ . Supongamos primero que el coste marginal de la educación es idéntico para los miembros de la misma familia ( $r_{ij} = r_i$ ) pero que la habilidad no tiene componente familiar (es decir,  $\text{cov}[b_{11}, b_{12}] = 0$ ). En este caso, todas las diferencias de educación dentro de las familias son debidas a diferencias de habilidad, mientras en la población en su conjunto solo una fracción  $f = \sigma_b^2 / (\sigma_b^2 + \sigma_r^2)$  de la varianza de la educación es atribuible a la habilidad. Como ya se mostró previamente, el componente de sesgo de endogeneidad en el estimador MCO de corte transversal es  $\psi_0 = k f$ . A partir de la ecuación (19) es fácil demostrar que  $\psi_{11} = k f / (1 - (1 - f)^2)$  y  $\psi_{12} = -k f (1 - f) / (1 - (1 - f)^2)$ . Por lo tanto,  $\psi_{11} - \psi_{12} = k$ , lo que implica que el estimador *intra-familia* tiene un sesgo de endogeneidad mayor que el estimador de corte transversal.

En el otro extremo, supongamos que la habilidad de los miembros de la misma familia es idéntica ( $b_{ij} = b_i$ ), pero que las preferencias no están correlacionadas dentro de las familias. En este caso, las diferencias de educación dentro de las familias se deben com-

<sup>48</sup> Si los supuestos de "efectos familiares puros" y de simetría son satisfechos, entonces  $\text{plim}(\tau_{11} - \tau_{21}) = \bar{\beta}$ .

<sup>49</sup> Asimismo, está estrechamente relacionado con el coeficiente de la diferencia de educación en un modelo de diferencias interfamiliares:  $\Delta \log y_i = \tau_\Delta S_i + \Delta e_i$ . Esta especificación es apropiada si las restricciones de simetría  $\lambda_{11} = \lambda_{22}$ ,  $\lambda_{12} = \lambda_{21}$ ,  $\psi_{11} = \psi_{22}$ ,  $\psi_{12} = \psi_{21}$  y  $\bar{S}_1 = \bar{S}_2$  son válidas, en cuyo caso  $\tau_{11} = \tau_{22}$  y  $\tau_{21} = \tau_{12}$ . Por ejemplo, en el caso de mellizos del mismo sexo la identidad de cada mellizo es arbitraria, por lo que un argumento de "intercambiability" sugiere que debería cumplirse la simetría. Bajo este supuesto  $\text{plim } \tau_\Delta = \text{plim}(\tau_{11} - \tau_{21}) = \text{plim}(\tau_{11} - \tau_{12})$ , a pesar de que la estimación de  $\tau_\Delta$  no es mecánicamente igual a la diferencia de las estimaciones de  $\tau_{11}$  y  $\tau_{12}$ .

pletamente a las diferencias en preferencias, incluso aunque en el conjunto de la población una proporción  $f$  de la varianza de la educación sea debida a diferencias de habilidad. Por lo tanto, el estimador *intra-familia* está libre de sesgo de endogeneidad, mientras el estimador MCO tiene un componente de sesgo de endogeneidad  $\psi_0 = k f$ . Así, las magnitudes relativas del sesgo de endogeneidad en los estimadores *intra-familia* y de corte transversal dependen de la contribución relativa de los diferenciales de habilidad a la varianzas de educación *intra-familia* y de corte transversal<sup>50</sup>. Un estimador *intra-familia* tendrá un sesgo menor si y solo si las diferencias de habilidad son determinantes menos importantes de la educación dentro de las familias que entre la población en su conjunto.

Los problemas de error de medida desempeñan un papel muy importante en la interpretación de las estimaciones de los modelos de hermanos y familias. Esto es particularmente cierto en estudios de gemelos idénticos, que tienden a tener niveles educativos altamente correlacionados. Por ejemplo, consideremos la estimación de la ecuación (17a) empleando medidas imprecisas de la educación para ambos gemelos. La fórmula del error de medida multivariante implica que el límite de probabilidad del coeficiente de la educación propia es:

$$\bar{\beta} \frac{R_0 - \rho^2}{1 - \rho^2} + (\lambda + \psi \bar{S}) \frac{1 - R_0}{1 - \rho^2} \times \frac{\text{cov}[S_{i1}, S_{i1} + S_{i2}]}{\text{var}[S_{i1}]},$$

donde  $R_0$  es la fiabilidad de la medida de educación y  $\rho$  es la correlación de la educación de los gemelos. Suponiendo que  $R_0 \approx 0,9$  y  $\rho \approx 0,75$  (véase por ejemplo, Ashenfelter y Rouse, 1998), esta fórmula supone que el límite de probabilidad del coeficiente de la educación propia es aproximadamente  $0,8 \bar{\beta} + 0,3 (\lambda + \psi \bar{S})$ .

Mucha de la literatura de gemelos se centra en la estimación de un modelo *intra-familia* en diferencias:

$$\Delta \log y_i = \tau_\Delta \Delta S_i + \Delta e_i.$$

Suponiendo que los supuestos de “efectos familiares puros” son satisfechos e ignorando errores de medida,

$$\text{plim } \tau_\Delta = \bar{\beta},$$

como se puede observar al diferenciar las ecuaciones (20a) y (20b). No obstante, el estimador *intra-familia* diferenciado es particularmente susceptible a errores de medida, dado que diferenciar dentro de las familias elimina gran parte de la variabilidad de la educación. En particular, si la fiabilidad de la educación observada es  $R_0$  y la correlación entre la educación de los miembros de la familia es  $\rho$ , entonces la fiabilidad de la diferencia de educación observada es

<sup>50</sup> Un argumento similar es aplicable a los sesgos asintóticos de los dos estimadores asociados con la correlación entre  $a_{ij}$  y  $S_{ij}$ .

$$R_{\Delta} = \frac{R_0 (1 - \rho)}{1 - \rho R_0}$$

Cuando  $R_0 \approx 0,9$  y  $\rho \approx 0,75$ , por ejemplo,  $R_{\Delta} \approx 0,7$ , lo que implica un sesgo de atenuación del 30% en la estimación MCO de  $\tau_{\Delta}$  para los gemelos idénticos. Entre los mellizos la correlación de la educación es menor: Ashenfelter y Krueger (1994) e Isacson (1997) estimaron una correlación entre mellizos de aproximadamente 0,55. Suponiendo  $R_0 \approx 0,9$  y  $\rho \approx 0,55$ ,  $R_{\Delta} \approx 0,8$ , así se esperaría un sesgo de atenuación del 20% en la estimación MCO de  $\tau_{\Delta}$  para los mellizos.

### ■ 1.3.8. Resumen

La tabla 1.3 resume algunos de los modelos, hipótesis, y ecuaciones estimadas clave que son útiles en la interpretación de la literatura sobre rendimientos educativos. Una estrategia de estimación no incluida en la tabla es la de variables instrumentales basadas en la comparación entre un grupo tratamiento cuasi experimental y un grupo comparación cuando el tratamiento tiene potencialmente diferentes efectos sobre el nivel educativo de diferentes subgrupos de población. Como ya se explicó anteriormente, bajo condiciones ideales ese estimador recuperará una media ponderada de los rendimientos marginales de la educación para diferentes subgrupos, donde la ponderación aplicada a cada subgrupo es la variación de educación inducida por el tratamiento. Esta media ponderada puede ser superior o inferior al rendimiento marginal medio de la educación, dependiendo del tipo de intervención y del grado de heterogeneidad de los rendimientos marginales.

Entre las implicaciones de los resultados resumidos en la tabla 1.3 están:

1. El estimador MCO tiene dos sesgos de habilidad respecto al rendimiento marginal medio de la educación ( $\bar{\beta}$ ): uno atribuible a la correlación entre educación y el intercepto de la función de ingresos ( $a$ ), y otro atribuible a la correlación entre educación y la pendiente de la función de ingresos ( $b$ ). Este último es claramente positivo, pero puede ser de pequeña magnitud si la heterogeneidad de los rendimientos de la educación es pequeña (o si las personas tienen una capacidad imperfecta para valorar su habilidad).
2. Las condiciones necesarias para que los estimadores por VI y función control ofrezcan una estimación consistente de  $\bar{\beta}$  en presencia de heterogeneidad en el rendimiento de la educación, son bastante estrictas. Fuentes plausibles de variación exógena en la elección educativa (tales como cambios en el coste de la educación) pueden no satisfacer estas condiciones, en cuyo caso las VI recuperarán la media ponderada de los rendimientos marginales para los subgrupos afectados.
3. Si el estimador MCO es sesgado al alza debido a habilidad no observada, se esperaría que un estimador VI basado en el *background* familiar estuviera incluso más sesgado al alza.

Tabla 1.3

**RESUMEN DE MODELOS, MÉTODOS DE ESTIMACIÓN Y LÍMITES DE PROBABILIDAD DE LOS ESTIMADORES**

MODELO	SUPUESTOS ADICIONALES	ECUACIÓN A ESTIMAR	LÍMITE DE PROBABILIDAD DEL ESTIMADOR
<b>I. Mínimos Cuadrados Ordinarios</b>			
$y_i = a_0 + b\bar{S}_i - 1/2k_1S_i^2 + a_i + (b_i - \bar{b})S_i$ $a_i = \lambda_0 (S_i - \bar{S}) + u_i$ $b_i = \bar{b} + \psi_0 (S_i - \bar{S}) + v_i$	$(a_i, b_i, S_i)$ Conjunta simétrica	$y_i = c + b_{OLS}S_i$	$b_{OLS}: \bar{\beta} + \lambda_0 + \psi_0\bar{S}, \bar{\beta} = \bar{b} - k_1\bar{S}$ con error de medida $b_{OLS}: R_0 (\bar{\beta} + \lambda_0 + \psi_0\bar{S})$ $R_0 = \text{confiabilidad de } S_i$
<b>Ila. Variables Instrumentales</b>			
$y_i = a_0 + b\bar{S}_i - 1/2k_1S_i^2 + a_i + (b_i - \bar{b})S_i$ $S_i = \pi Z_i + \xi_i$	(a) $E[\eta_i Z_i] = 0$ (b) $E[a_i Z_i] = 0$ (c) $E[(b_i - \bar{b}) Z_i] = 0$ (d) $E[(b_i - \bar{b})^2 Z_i] = \sigma_b^2$ (e) $E[(\xi_i b_i] = \rho_1(b_i - \bar{b})$	$S_i = \pi Z_i + \xi_i$ $y_i = c + b_{IV}S_i$	$b_{IV}: \bar{\beta}$
<b>Iib. Función Control</b>			
Igual que Ila	(a) - (c) anterior más (d) $E[a_i S_i Z_i] = \lambda_1 S_i + \lambda_2 Z_i$ (e) $E[b_i S_i Z_i] = \psi_1 S_i + \psi_2 Z_i$	$S_i = \pi Z_i + \xi_i$ $y_i = c + b_{CF} S_i + e_0 \xi_i + e_1 S_i \xi_i$	$b_{CF}: \bar{\beta}$
<b>III. Modelos de Background Familiar</b>			
$y_i = a_0 + b\bar{S}_i - 1/2k_1S_i^2 + a_i + (b_i - \bar{b})S_i$ $a_i = \lambda_1 (S_i - \bar{S}) + \lambda_2 (F_i - \bar{F}) + u_i$ $b_i = \bar{b} + \psi_1 (S_i - \bar{S}) + \psi_2 (S_i - \bar{S}) + v_i$ $\rho = \text{correlación } (S_i, F_i)$	$(a_i, b_i, S_i, F_i)$ Conjunta simétrica	$y_i = c + b_{OIV} + S_i + gF_i$ $S_i = \pi F_i + e_i$ $y_i = c + b_{IV} \bar{S}_i$	$b_{OIV}: \bar{\beta} + \lambda_1 + \psi_2 \bar{S}$ $g = \lambda_2 + \psi_2 \bar{S}$ $b_{IV}: \bar{\beta} + \lambda_1 + \psi_1 \bar{S} + (\lambda_2 + \psi_2 \bar{S})/\pi_0$ Con error de medida $b_{OIV}: R_1 (\bar{\beta} + \lambda_1 + \psi_1 \bar{S}), R_1 = (R_0 - \rho)/(1 - \rho)$
<b>III. Modelos de hermanos/gemelos</b>			
$y_{ij} = a_0 + b\bar{S}_{ij} - 1/2k_1S_{ij}^2 + a_{ij} + (b_{ij} - \bar{b})S_{ij}$ $a_{ij} = \lambda_{j1} (S_{ij} - \bar{S}_1) + \lambda_{j2} (S_{ij} - \bar{S}_2) + u_{ij}$ $b_{ij} = \bar{b} + \psi_{j1} (S_{ij} - \bar{S}_1) + \psi_{j2} (S_{ij} - \bar{S}_2) + v_{ij}$	$(a_{ij}, b_{ij}, S_{i1}, S_{i2})$ Conjunta simétrica Anterior más $\lambda_{12} = \lambda_{21}; \psi_{12} = \psi_{21}$ ; $\lambda_{11} = \lambda_{22}; \psi_{11} = \psi_{22}$ ; $\bar{S}_1 = \bar{S}_2 = \bar{S}$ $j = 1, 2$ "intercambiable"	$y_{i1} = c_1 + \tau_{11}S_{i1} + \tau_{12}S_{i2}$ $y_{i2} = c_2 + \tau_{21}S_{i1} + \tau_{22}S_{i2}$ Igual que antes o $\Delta y_i = \tau_{\Delta} \Delta S_i$	$\tau_{11}: \bar{\beta} + \lambda_{11} + \psi_{11}\bar{S}_1 + \tau_{12}: \lambda_{12} + \psi_{12}\bar{S}_1$ $\tau_{22}: \bar{\beta} + \lambda_{22} + \psi_{22}\bar{S}_2 + \tau_{21}: \lambda_{21} + \psi_{21}\bar{S}_2$ $\tau_{11}: \tau_{22}: \bar{\beta} + \lambda_{11} + \psi_{11}\bar{S}$ $\tau_{12}: \tau_{21}: \lambda_{12} + \psi_{12}\bar{S}$ $\tau_{\Delta}: \bar{\beta} + \lambda_{11} - \lambda_{12} + (\psi_{11} - \psi_{12})\bar{S}$ $\tau_{\Delta}: R_{\Delta} (\bar{\beta} + \lambda_{11} - \lambda_{12} + (\psi_{11} - \psi_{12})\bar{S})$ Con error de medida $R_{\Delta} = \text{confiabilidad de } \Delta S_i$

4. Si gemelos o hermanos tienen habilidades *idénticas* (y la distribución de habilidad entre gemelos es la misma que entre la población en su conjunto) entonces un estimador *intra-familia* obtendrá una estimación asintóticamente insesgada del rendimiento marginal medio de la educación. De otro modo, el estimador *intra-familia* puede estar más o menos sesgado que el correspondiente estimador MCO de corte transversal por el efecto de la habilidad no observada, dependiendo de la proporción relativa de la varianza de la educación atribuible a diferencias de habilidad dentro de las familias frente a las diferencias en el conjunto de la población.
5. Los sesgos por errores de medida son potencialmente importantes al interpretar las estimaciones por diferentes procedimientos. Las estimaciones convencionales por MCO están probablemente sesgadas a la baja en un 10%; las estimaciones MCO que controlan por *background* familiar (o por la educación de un hermano) pueden estar sesgadas a la baja en un 15% o más; y estimaciones *intra-familia* diferenciadas, pueden estar sesgadas a la baja entre un 20 y un 30%, probablemente siendo el rango superior más probable en el caso de utilizar gemelos idénticos.

#### ■ 1.4. UNA REVISIÓN SELECTIVA DE ESTUDIOS EMPÍRICOS RECIENTES

Ahora paso a presentar una revisión selectiva de la literatura reciente sobre la estimación de los rendimientos educativos. La revisión se agrupa en tres conjuntos de resultados: estimaciones del rendimiento de la educación por variables instrumentales basadas en aspectos institucionales del sistema educativo; estimaciones basadas en controlar el *background* familiar o bien que usan el *background* familiar como un instrumento de la educación; y estimaciones basadas en la educación e ingresos de gemelos. También reviso brevemente los recientes esfuerzos para modelizar la heterogeneidad observable en los rendimientos educativos. Una corriente de la literatura que no incluyo son los estudios de rendimientos de la educación que intentan controlar la habilidad empleando resultados observados de tests. Algunas de las sutiles cuestiones que implica el desarrollo de un enfoque causal para la interpretación de las puntuaciones de las pruebas, los niveles de educación, e ingresos, son considerados por Griliches (1977, 1979), Chamberlain (1977) y Chamberlain y Griliches (1975, 1977).

##### ■ 1.4.1. Variables instrumentales basadas en aspectos institucionales del sistema educativo

Uno de los nuevos enfoques de investigación más importante en la literatura reciente sobre educación, es el uso de aspectos institucionales del sistema educativo como una fuente de información identificadora creíble para esclarecer los efectos causales de la edu-

cación<sup>51</sup>. La tabla 1.4 resume siete estudios recientes que estiman el rendimiento de la educación empleando variables instrumentales, basados en esta idea. Para cada estudio se ofrece la estimación MCO y VI obtenida de la misma muestra y con las mismas variables de control.

El artículo de referencia de Angrist y Krueger (1991) usa el trimestre de nacimiento (interactuando con el año de nacimiento o el estado de nacimiento en algunas especificaciones) como instrumento de la educación. Los autores muestran que los hombres nacidos entre 1930 y 1959 al principio del año, tienen ligeramente menos educación que los hombres nacidos en una fecha más tardía dentro del año —efecto que atribuyen a las leyes de educación obligatoria. Angrist y Krueger observan que las personas nacidas en el mismo año natural normalmente empiezan a estudiar al mismo tiempo. Como resultado, los individuos nacidos al principio del año alcanzan la edad mínima obligatoria en un nivel menor que las personas nacidas posteriormente durante el año, permitiendo que aquellos que quieren abandonar tan pronto como sea legalmente posible, dejen la escuela con menor educación. Suponiendo que el trimestre de nacimiento es independiente de los factores de preferencias y habilidad, este fenómeno genera variación exógena en la educación que se puede usar en la estimación por VI. Es importante resaltar que las leyes de educación obligatoria presumiblemente aumentan la educación de las personas que en otro caso escogerían bajos niveles de educación. Si estos individuos tienen mayor o menor rendimiento marginal de la educación que otras personas, un estimador VI basado en el trimestre de nacimiento, puede sobre- o sub- estimar el rendimiento marginal medio de la educación de toda la población.

El análisis empírico de Angrist y Krueger confirma que el patrón trimestral en el nivel educativo se comporta de forma paralela al patrón de ingresos. Como se muestra en la tabla 1.4, las estimaciones por VI del rendimiento de la educación son generalmente mayores que las correspondientes estimaciones por MCO, aunque para algunas cohortes y especificaciones los dos estimadores son muy similares, y en ningún caso la diferencia entre los estimadores VI y MCO es estadísticamente significativa.

Los hallazgos de Angrist y Krueger han generado mucho interés y algunas críticas. Bound *et al.* (1995) señalan que varios de los modelos de VI de Angrist y Krueger (específicamente, aquellos que usan interacciones entre el trimestre y el estado de nacimiento como predictores de la educación) incluyen gran número de instrumentos débiles, y están por tanto, asintóticamente sesgados hacia las correspondientes estimaciones MCO. Este sesgo de “instrumentos débiles” no es un problema para las especificaciones presentadas en la tabla 1.4, que se basan en un conjunto más parsimonioso de instrumentos. Adicionalmente, en la medida que las estimaciones VI de Angrist y Krueger están *por encima* de las correspon-

---

<sup>51</sup> Esta idea también está demostrando su utilidad en los estudios sobre el efecto de la calidad de la educación. Por ejemplo, Angrist y Lavy (1997) utilizan la información sobre el tamaño máximo de las clases para identificar el efecto del tamaño de clase en el aprovechamiento estudiantil.

Tabla 1.4

## ESTIMACIONES MCO Y VI DEL RENDIMIENTO DE LA EDUCACIÓN CON INSTRUMENTOS BASADOS EN LAS CARACTERÍSTICAS DEL SISTEMA EDUCATIVO\*

AUTOR	MUESTRA E INSTRUMENTOS	COEFICIENTES DE EDUCACIÓN		
		MCO	VI	
1. Angrist y Krueger (1991)	Datos del Censo de 1970 y 1980, Hombres.	1920-1929 cohorte en 1970	0,070	0,101
	Instrumentos son el trimestre de nacimiento interactuando con el año de nacimiento.	1930-1939 cohorte en 1980	(0,000)	(0,033)
	Controles incluye cuadrático en la edad e indicadores de raza.	1940-1949 cohorte en 1980	0,063	0,060
	Estado civil, residencia urbana		(0,000)	(0,030)
			0,052	0,078
		(0,000)	(0,030)	
2. Staiger y Stock (1997)	Censo 1980, Hombres. Instrumentos son el trimestre de nacimiento interactuando con estado y año de nacimiento. Controles son los mismos que en Angrist y Krueger, más indicadores por estado de nacimiento.	1930-1939 cohorte en 1980	0,063	0,098
	Estimaciones Máximo Verosímil con Información Limitada	1940-1949 cohorte en 1980	(0,000)	(0,015)
			0,052	0,088
			(0,000)	(0,018)
3. Kane y Rouse (1993)	NLS, Clase de 1972, mujeres.	Modelos sin puntajes de pruebas o educación de los padres	0,080	0,091
	Instrumentos son matrículas en universidades estatales de 2 y 4 años y distancia a la universidad más cercana. Controles incluyen raza, estatus tiempo parcial, experiencia.	Modelos con puntajes de pruebas y educación de los padres	(0,005)	(0,033)
	Nota: Educación medida en unidades de créditos universitarios equivalentes		0,063	0,094
			(0,005)	(0,042)
4. Card (1995b)	NLS, Hombres jóvenes (Cohorte 1966)	Modelos que usan proximidad a universidad como instrumento (ingresos 1976)	0,073	0,132
	Instrumento indicador de cercanía a universidades de 4 años en 1966, o la interacción de estos con la educación de los padres	Modelos que usan proximidad a universidad x <i>background</i> familiar como instrumento	(0,004)	(0,049)
	Controles incluye raza, experiencia (tratadas como endógenas), región, y educación de los padres		—	0,097
				(0,048)
5. Conneely y Uusitalo (1997)	Hombres Finlandeses que sirven en la Armada en 1982, y estuvieron trabajando Tiempo completo en trabajos civiles en 1994.	Modelos que excluyen educación de los padres e ingresos	0,085	0,110
	Datos Administrativos de ingresos y educación. Instrumento es vivir en ciudad Universitaria en 1980. Controles incluye cuadrático en experiencia e Ingresos y educación de los padres.	Modelos que incluyen educación de los padres e ingresos	(0,001)	(0,024)
			0,083	0,098
		(0,001)	(0,035)	

Tabla 1.4 (continuación)

### ESTIMACIONES MCO Y VI DEL RENDIMIENTO DE LA EDUCACIÓN CON INSTRUMENTOS BASADOS EN LAS CARACTERÍSTICAS DEL SISTEMA EDUCATIVO\*

AUTOR	MUESTRA E INSTRUMENTOS	COEFICIENTES DE EDUCACIÓN	
		MCO	VI
6. Maluccio (1997)	Encuesta Multipropósito Bicol (rural Filipinas): Modelos que no controlan por	0,073	0,145
	hombres y mujeres asalariados selección de su situación laboral o localización	(0,011)	(0,041)
	Entre 20-44 años en 1994, cuyas familias fueron entrevistadas en 1978. Modelos con corrección por	0,063	0,113
	Instrumentos son distancia escuela selección localización o situación secundaria más cercana e indicador para laboral escuela secundaria local privada. Controles incluye cuadrático en edad.	(0,006)	(0,033)
7. Harmon y Walker (1995)	Encuesta de presupuestos familiares Británica 1978-1986 (hombres). Instrumentos son indicadores para cambios en edad mínima de escolarización en 1947 y 1973. Controles incluye cuadrático en edad, año, encuesta y región y región	0,061 (0,001)	0,153 (0,015)

Nota: Véase el texto de las fuentes y la información sobre los estudios individuales.

dientes estimaciones MCO, se podría inferir que las estimaciones asintóticamente insesgadas del efecto causal de la educación son incluso superiores. Esto es confirmado por los hallazgos de Staiger y Stock (1997), quienes re-analizan las muestras del Censo de 1980 usada por Angrist y Krueger y calculan diversos intervalos de confianza con validez asintótica para estimaciones estándar por VI y estimaciones Máximo Verosímil con Información Limitada (MVIL). Las estimaciones MVIL preferidas por Staiger y Stock, empleando el trimestre de nacimiento interactuando con el estado y el año de nacimiento como instrumentos, son presentadas en la fila 2 de la tabla 1.4. Estas estimaciones son algo superiores a las correspondientes estimaciones VI convencionales, y entre un 50 y un 70% mayores que las estimaciones MCO.

Una segunda crítica a los hallazgos de Angrist y Krueger, planteada por Bound y Jaeger (1996), es que el trimestre de nacimiento puede estar correlacionado con diferencias en habilidad no observada. Bound y Jaeger evalúan el nivel de educación de cohortes de hombres que no estuvieron sujetos a las instituciones de educación obligatoria, y hallan alguna evidencia de patrón estacional. También discuten la evidencia de la literatura de sociobiología y psicobiología que sugiere que el trimestre de nacimiento está relacionado con el *background* familiar y la incidencia de enfermedades mentales.

Para evaluar las diferencias en el *background* familiar por el trimestre de nacimiento para cohortes aproximadamente comparables con las del estudio de Angrist y Krueger, comparé los niveles medios de educación de los padres por trimestre de nacimiento para niños menores de un año de edad en el Censo de 1940<sup>52</sup>. Los años medios de educación de las madres de niños nacidos entre los trimestres I, II, III y IV, son 9,04, 8,95, 8,97, y 8,95, respectivamente (con errores estándar alrededor de 0,05). La correspondiente media de la educación de los padres son 8,61, 8,50, 8,52, y 8,58. Estas comparaciones no ofrecen indicios de que los niños nacidos en el primer trimestre procedan de familias con *background* familiar relativamente desfavorecido, y sugieren que los patrones estacionales identificados por Angrist y Krueger son probablemente ocasionados por diferencias en el *background* familiar.

El tercer estudio resumido en la tabla 1.4, de Kane y Rouse (1993), se centra en la valoración relativa del mercado laboral de los créditos de universidades de 4 años (regular college) o 2 años (junior college). Sus resultados sugieren que los créditos concedidos por los dos tipos de universidades son intercambiables: a la luz de esta conclusión los autores miden la educación en términos del total de créditos universitarios equivalentes. En el análisis del efecto de los créditos universitarios en los ingresos, Kane y Rouse comparan las especificaciones MCO contra los modelos por VI que emplean como instrumentos la distancia a la universidad de 2-años o 4-años más cercana y las tasas de matrícula específicas de estado. Sus estimaciones por VI basadas en estos instrumentos son un 15-50% superiores a la correspondiente especificación MCO.

Dos estudios posteriores, de Card (1995b) y Conneely y Uusitalo (1997), evalúan la educación y los diferenciales de ingreso, asociados con crecer cerca de una universidad. El estudio de Card (1995b) encuentra que cuando la proximidad a la universidad es empleada como instrumento de la educación en la muestra de hombres jóvenes del *National Longitudinal Survey* (NLS), el estimador VI resultante es considerablemente superior al correspondiente estimador MCO, aunque más impreciso. De acuerdo con la idea de que la accesibilidad importa más para individuos en el margen de continuar con su educación, se encuentra que la proximidad a la universidad tiene un efecto mayor para niños de padres menos educados. Esto sugiere una especificación alternativa que utiliza las interacciones de la proximidad a la universidad con variables de *background* familiar como instrumentos de la educación, e incluye la proximidad a la universidad como una variable de control directo. La estimación por VI de esta especificación interactuada es algo menor que la estimación que emplea sólo la proximidad a la universidad, pero es aún alrededor de un 30% superior a la estimación MCO.

El estudio de Conneely y Uusitalo (1997) emplea una base de datos de Finlandia muy rica que combina información del *background* familiar, puntuaciones de tests militares y

<sup>52</sup> La información sobre el trimestre de nacimiento sólo está disponible en el Censo de 1940 para los niños menores de 1 año de edad. Hay 19.089 niños menores de 1 año en el archivo de uso público, de los cuales el 98,4% se puede emparejar a una mujer cabeza de familia y el 95,3% se puede emparejar a un hombre cabeza de familia.

datos administrativos de ingresos, para hombres que sirvieron en las fuerzas armadas en el año 1982. Al igual que Kane y Rouse (1993) y Card (1995b), estos autores hallan que las estimaciones por VI de los rendimientos de la educación basados en la proximidad a las universidades, superan las correspondientes estimaciones MCO entre un 20-30%, dependiendo de que otros controles se añaden en el modelo. Cabe señalar que estos tres estudios presentan modelos que controlan por un grupo bastante detallado de características de *background* familiar. Estos controles son deseables si las familias que viven cerca de universidades tienen diferente *background* familiar, y si el *background* familiar tiene algún efecto causal independiente sobre los ingresos. Las estimaciones por VI de Conneely y Uusitalo controlando la educación y los ingresos de los padres, es inferior a la estimación por VI que excluye estos controles, pero es aún superior a las estimaciones más sencillas por MCO sin controles por *background* familiar. A pesar del tamaño bastante grande de su muestra (cerca de 22.000 observaciones) y de la elevada calidad de sus datos, las estimaciones por VI de Conneely y Uusitalo son algo imprecisas y no significativamente diferentes de sus estimaciones MCO<sup>53</sup>.

El sexto estudio en la tabla 1.4, de Maluccio (1997), aplica la idea de proximidad a la escuela a datos rurales de Filipinas. Maluccio combina la información de educación e ingresos para una muestra de adultos jóvenes, con datos de los hogares de sus padres, incluyendo la distancia a la escuela secundaria más cercana y un indicador para la presencia de una escuela secundaria privada local. Estas variables tienen un efecto relativamente fuerte sobre la educación en esta muestra. Maluccio estima por MCO y modelos convencionales por VI, empleando la proximidad a la escuela como un instrumento, así como los modelos por VI que incluyen corrección por selección de situación laboral y localización. Ambas estimaciones por VI son sustancialmente superiores a las correspondientes estimaciones MCO. El análisis de Maluccio sugiere que la fiabilidad de su variable de educación es algo menor que las bases de datos convencionales estadounidenses o europeas ( $R_0 \approx 0,8$ ), lo que explica algo de las diferencias entre las estimaciones MCO y por VI. Desafortunadamente, Maluccio no presenta modelos MCO y VI que controlen por *background* familiar. Presenta modelos VI que utilizan la educación y riqueza de los padres como instrumentos adicionales para la educación, conduciendo a unas estimaciones por VI ligeramente menores, pero algo más precisas.

El último estudio incluido en la tabla 1.4, de Harmon y Walker (1995), examina los rendimientos de la educación entre una muestra relativamente grande de cabezas de familia varones británicos. Harmon y Walker usan como variables instrumentales de la educación un par de variables *dummies* que recogen cambios en la edad mínima obligatoria de escolarización en Gran Bretaña —de 14 a 15 en 1947, y de 15 a 16 en 1973. Éstas son efectivamente *dummies* de cohorte que distinguen entre hombres nacidos antes de 1932, de aquellos nacidos entre 1933 a 1957 y aquellos nacidos después de 1957. Como se mues-

<sup>53</sup> Conneely y Uusitalo también implementan un estimador de la función control más general, como el que se ha descrito en páginas anteriores.

tra en la tabla 1.4 su estimación por VI es considerablemente superior a la MCO (2,5 veces mayor), y relativamente precisa. Existen varios aspectos de su estrategia de estimación que sugieren, no obstante, la necesidad de tener precaución en la interpretación de estos resultados. Muy importante, el cambio de la ley de 1947 —que es la fuente más importante de identificación en sus resultados— viene justo después de la Segunda Guerra Mundial<sup>54</sup>. Además, Harmon y Walker no permiten el crecimiento sistemático en el nivel educativo para cohortes consecutivas de hombres, más allá del atribuible a los cambios legislativos en 1947 y 1973<sup>55</sup>. Ambos factores puede sesgar su estimador VI al alza.

Adicionalmente a los estudios incluidos en la tabla 1.4, otros estudios recientes han empleado técnicas de VI para estimar el rendimiento de la educación. Un ejemplo innovador es Hausman y Taylor (1981), que usa las medias de tres variables cambiantes en el tiempo (edad, indicadores de la incidencia de problemas de salud y desempleo) como instrumentos de la educación en un modelo de panel de datos de los ingresos obtenidos por hombres *prime-age*. Hausman y Taylor encuentran que el rendimiento de la educación aumenta desde cerca de 0,07 en las especificaciones MCO a 0,12-0,13 en sus especificaciones por VI. A pesar de que estudios más recientes no han seguido directamente la metodología de Hausman y Taylor, su uso de la edad media como instrumento de la educación es equivalente a emplear una variable lineal de cohorte, y es por lo tanto similar en esencia al estudio de Harmon y Walker.

Un muy reciente estudio de Ichino y Winter-Ebmer (1998), también utiliza la cohorte de nacimiento como fuente de variación en el nivel educativo. En particular, Ichino y Winter-Ebmer se centran en los ingresos y educación de los varones austriacos y alemanes, nacidos entre 1930 y 1935. Los autores argumentan que la Segunda Guerra Mundial tuvo un efecto particularmente fuerte sobre el nivel educativo de los niños que alcanzaron la adolescencia durante la guerra y vivieron en países directamente afectados por las hostilidades. Empleando datos para 14 países hallan diferencias relativamente grandes en la educación alcanzada por los niños en las cohortes 1930-1935 en los países que fueron más duramente afectados por la guerra (por ejemplo, Alemania, Austria y el Reino Unido) pero diferencias relativamente pequeñas para estas cohortes en otros países (por ejemplo, EE.UU. e Irlanda). Cuando emplean un indicador para la cohorte 1930-1935 como instrumento de nivel educativo bajo, encuentran que la desventaja en ingresos aproximadamente dobla su valor MCO. Aunque se podría pensar que la cohorte 1930-1935 sufrió otras desventajas además de la interrupción de su carrera educativa, estos resultados son comparables con los de Harmon y Walker (1995) en términos de la magnitud de la diferencia VI/MCO.

<sup>54</sup> Ichino y Winter-Ebmer (1998), documentan que en toda Europa, el nivel educativo de los niños nacidos entre 1930 y 1935 fue considerablemente inferior al de los niños nacidos justo antes o después.

<sup>55</sup> Sus especificaciones controlan por edad y año de la encuesta. Se puede inferir la presencia de importantes efectos de cohorte del hecho de que su efecto año de la encuesta muestra un aumento de 0,5 año en el nivel educativo entre las encuestas de 1979 y 1986, controlando por edad y los indicadores de educación obligatoria.

Otro estudio no presentado en la tabla 1.4, de Angrist y Krueger (1992), examina el efecto potencial del comportamiento “evasor del reclutamiento obligatorio” sobre la educación y los ingresos de hombres que estaban en riesgo de alistamiento en el sorteo del servicio militar de 1970-1973 para la guerra de Vietnam. Dado que los estudiantes matriculados podían obtener excepciones a la llamada a filas, muchos investigadores argumentan que el sorteo de alistamiento llevó a mayores tasas de matrícula en la universidad, particularmente para hombres cuyo número de sorteo implicaba el mayor riesgo de alistamiento. Si esto es cierto, se podrían emplear los números del sorteo de alistamiento —que fueron aleatoriamente asignados por día de nacimiento— como instrumento de la educación. Aunque Angrist y Krueger (1992) presentan estimaciones por VI basadas en esta idea, un estudio posterior (Angrist y Krueger, 1995), mostró que la relación entre los números del sorteo y la educación alcanzada es muy débil. De hecho, las diferencias de educación entre grupos de hombres con diferentes números de sorteo no son estadísticamente significativas. Así, las estimaciones por VI están sujetas a la crítica de instrumentos débiles de Bound *et al.* (1995), y son esencialmente poco informativas sobre el efecto causal de la educación<sup>56</sup>.

Una conclusión que surge de los resultados de la tabla 1.4 y de otros estudios basados en VI, es que las estimaciones del rendimiento de la educación por variables instrumentales generalmente exceden la correspondiente estimación MCO —habitualmente en un 30% o más. Si se supone a priori que los métodos MCO llevan a estimaciones sesgadas al alza, las estimaciones por VI de muchos estudios recientes con aún más elevados resultados nos sitúan ante una paradoja. Se han desarrollado una serie de hipótesis para explicar la paradoja. La primera —sugerida por Bound y Jaeger (1996), por ejemplo— es que las estimaciones por VI están *incluso más* sesgadas al alza que las correspondientes estimaciones MCO, por diferencias no observadas entre las características de los grupos tratamiento y comparación, implícitas en el enfoque de VI. Sin la menor duda, esta es una explicación plausible de parte de la diferencia entre MCO y VI en estudios que no controlan directamente el *background* familiar, pero es menos creíble para estudios que incluyen controles de *background* familiar.

Una segunda explicación —propuesta por Griliches (1977) y retomada por Angrist y Krueger (1991)— es que los sesgos de habilidad en las estimaciones MCO del rendimiento de la educación son relativamente pequeños, y que las diferencias entre las estimaciones VI y MCO de la tabla 1.4 reflejan el sesgo a la baja de las estimaciones MCO atribuibles a errores de medida. La imprecisión de la mayoría de las estimaciones por VI de la tabla 1.4 hace difícil descartar esta explicación analizando cada uno de los estudios. No obstante, dado que el sesgo de errores de medida sólo puede explicar por sí mismo un

<sup>56</sup> Las estimaciones VI convencionales son generalmente iguales o ligeramente superiores a las estimaciones MCO. Angrist y Krueger (1995) proponen un método de VI por “división de la muestra”, para tratar el problema de la debilidad de los instrumentos. Las estimaciones por VI con muestra dividida son todas muy imprecisas.

diferencial del 10% entre MCO y VI, parece poco probable que tantos estudios hallaran grandes diferenciales positivos entre las estimaciones VI y MCO sólo debidos a errores de medida<sup>57</sup>.

Una tercera posibilidad, sugerida por Ashenfelter y Harmon (1998) en una reciente revisión de la literatura sobre rendimientos de la educación, es el argumento del “sesgo de publicación”. Los autores sugieren que en una búsqueda entre especificaciones alternativas para las estimaciones por VI estadísticamente significativas, es más probable que un investigador seleccione una especificación que ofrezca una estimación elevada del rendimiento educativo. Como evidencia de este comportamiento presentan una correlación positiva entre diferentes estudios del diferencial VI-MCO en los rendimientos estimados y el error muestral de la estimación por VI<sup>58</sup>.

Aunque estos tres argumentos son sugerentes, considero que una cuarta explicación, basada en la heterogeneidad subyacente en los rendimientos de la educación es también potencialmente importante. Factores como la educación obligatoria o la accesibilidad a las escuelas es más probable que afecten las elecciones educativas de los individuos que de otro modo tendrían un nivel educativo relativamente bajo. Si estos individuos tienen un rendimiento marginal de la educación superior a la media, entonces se podría esperar que los estimadores de variables instrumentales basados en la educación obligatoria o en la proximidad a la escuela ofrezcan estimaciones de los rendimientos de la educación superiores a las correspondientes estimaciones MCO. Una condición necesaria para este fenómeno es que el rendimiento marginal de la educación esté negativamente correlacionado con el nivel de educación de la población. En el modelo presentado en la sección 3, la covarianza del rendimiento de la educación es  $E[\beta_i(S_i - \bar{S})] = (kf - k_1)\text{Var}[S_i]$ , donde  $k = k_1 + k_2$  y  $f$  es la parte de la varianza del nivel de educación atribuible a la variación en la habilidad. Si las tasas de descuento individuales son constantes (es decir,  $k_2 = 0$ ) esta covarianza es necesariamente negativa. Incluso si los individuos tienen tasas marginales de descuento crecientes (por factores de preferencias o restricciones financieras), los rendimientos marginales de la educación serán mayores para individuos menos educados si las diferencias de habilidad no son “demasiado importantes” en la determinación del nivel de educación, y si el rendimiento marginal de la educación es decreciente. En este caso, las estimaciones por VI de los rendimientos de la educación basados en cambios institucionales que aumentan los niveles de educación de los subgrupos menos educados pueden ser superiores a la correspondientes estimaciones MCO.

<sup>57</sup> Una salvedad a esta conclusión es la posibilidad de que los errores de medida sean mayores, o estén más sistemáticamente correlacionados con los niveles de educación, para los individuos más afectados por las intervenciones sobre las que se basan los análisis de la tabla 4. Kane *et al.* (1997) encuentran alguna evidencia al respecto.

<sup>58</sup> En los estudios incluidos en la tabla 4, el diferencial VI-MCO está negativamente relacionado con el error muestral de la estimación VI, aunque la correlación es positiva si se excluye el estudio de Harmon-Walker.

#### ■ 1.4.2. Estimaciones empleando *background* familiar como control o como instrumento

La tabla 1.5 resume algunos hallazgos sobre el uso del *background* familiar (generalmente, la educación parental o de un hermano) como variable control o como instrumento en modelos de rendimiento de la educación. Para la mayoría de estudios presentados en la tabla, presento tres estimaciones del rendimiento educativo: una estimación MCO que excluye controles de *background* familiar; una estimación MCO que introduce una o más características de *background* familiar y un estimador VI que utiliza *la(s) misma(s)* variable(s) de *background* familiar como instrumento para la educación. Para dos de estos estudios (Miller *et al.*, 1995; Ashenfelter y Rouse, 1998), también presento estimaciones por VI corregidas por error de medida para modelos que incluyen un indicador de la educación del individuo y un indicador de la educación de su hermano como instrumentos<sup>59</sup>. Se debe tener en cuenta que la mayoría de los estudios descritos en la tabla 1.5 no se centran directamente en las especificaciones que se han presentado, sino que presentan estos resultados de forma adicional.

El primer grupo de estudios en la tabla emplean la educación parental como indicador de *background* familiar. Los estudios de Card (1995b) y Conneely y Uusitalo (1997) ya han sido descritos<sup>60</sup>. Preparé específicamente para esta revisión las estimaciones con la muestra del *General Social Survey* (GSS)<sup>61</sup>. El artículo de Ashenfelter y Zimmerman (1997) utiliza la educación del padre como variable de *background* familiar en un grupo de sus modelos, y la educación del hermano en otro. Con la excepción de los resultados para mujeres en el GSS, los resultados para estos cuatro estudios son notablemente consistentes. En los cuatro casos la inclusión de la educación de los padres como variable control (o conjunto de controles) disminuye rendimiento de la educación en un 5-10% (aproximadamente la magnitud de la reducción esperada como resultado de errores de medida); mientras que el uso de la educación de los padres como instrumento lleva a estimaciones VI que son al menos un 15% superiores a las correspondientes estimaciones por MCO. Adicionalmente (aunque no se presenta en la tabla), el coeficiente de la variable de educación parental es positivo y significativo pero de pequeña magnitud. Para las mujeres en la muestra del GSS la inclusión de la educación de la madre prácticamente no tiene efecto sobre el rendimiento de la educación de las mujeres, y la mayor educación de la madre tiene un efecto negativo muy pequeño sobre los ingresos. Como consecuencia la estimación por VI para la muestra femenina del GSS es ligeramente menor que la estimación MCO<sup>62</sup>.

<sup>59</sup> En concreto, siguiendo el ejemplo de Ashenfelter y Krueger (1994), Miller *et al.*, y Ashenfelter y Rouse hacen uso de la información recopilada sobre la educación de gemelos y la de sus hermanos.

<sup>60</sup> La estimación por VI asociada con los datos de mi estudio de 1995b no se incluye en la versión publicada del documento.

<sup>61</sup> La GSS tiene la ventaja de incluir una amplia muestra de hombres y mujeres. La información de ingresos se basa en ingresos anuales: se imputó intervalos para los puntos medios de los datos categóricos de la encuesta.

<sup>62</sup> Recordemos que estos tres estimadores están mecánicamente relacionados. Si la educación de la madre tiene un efecto negativo al controlar la educación de la hija, entonces la estimación por VI, con educación de la madre como instrumento, es necesariamente menor que la estimación MCO.

Tabla 1.5

## LAS ESTIMACIONES DEL RENDIMIENTO DE LA EDUCACIÓN CON Y SIN EL CONTROL DE LOS ANTECEDENTES FAMILIARES, Y VI ESTIMACIONES A PARTIR DE LOS ANTECEDENTES FAMILIARES

AUTOR	MUESTRA Y VARIABLE(S) DE <i>BACKGROUND</i> FAMILIAR	COEFICIENTES MCO		VI
		NO CONTROL	CONTROL	COEFICIENTE
1. Card (1995b)	NLS, Hombres jóvenes (véase cuadro 4). Variables <i>background</i> familiar son educación ambos padres (efecto principal e interacciones) más estructura familiar <sup>b</sup>	0,073 (0,006)	0,069 (0,006)	0,084 (0,009)
2. Este capítulo	GSS, de adultos cabeza de familia entre 24-61 años en 1974-1996. Ingresos anuales (imputados de datos categóricos). Controles incluye cúbico en edad, raza, año de la encuesta y región. Variable <i>background</i> familiar educación madre	Hombres (N = 7.860) 0,073 (0,003)	0,067 (0,003)	0,106 (0,007)
		Mujeres (N = 7.500) 0,112 (0,004)	0,113 (0,004)	0,110 (0,011)
3. Conneely y Uusitalo (1997)	Hombres Finlandia veteranos (véase cuadro 4) variable <i>background</i> familiar es educación de los padres	0,085 (0,001)	0,082 (0,001)	0,114 (0,006)
4. Ashenfelter y Zimmerman (1997)	NLS hombres jóvenes (Cohorte 1966) fusionada con NLS, hombres mayores. variables <i>background</i> familiar son Educación de los padres o hermanos. Controles incluye cuadrático en edad	Hermano 1, utilizando educación otro hermano 0,059 (0,014)	0,052 (0,015)	0,080 (0,027)
		Hijos utilizando educación de padre 0,057 (0,009)	0,049 (0,009)	0,109 (0,025)
5. Miller <i>et al.</i> (1995)	Registro Australiano gemelos (hombres y mujeres gemelos idénticos). Ingreso imputado de ocupación. Variable <i>background</i> familiar educación gemelo.	No permitiendo errores medición 0,064 (0,002)	0,048 (0,003)	—
6. Ashenfelter y Rouse (1998)	Controles incluye cuadrático en edad y estado civil 1991-1993 Encuesta Gemelos Princeton (hombres y mujeres). Gemelos Idénticos. Variable <i>background</i> familiar es educación gemelo. Controles incluye cuadrático en edad, género y raza	VI usa reporte gemelos <sup>c</sup> 0,073 (0,003)	0,078 (0,009)	—
		No permitiendo errores de medición 0,102 (0,010)	0,092 —	
		VI usa reporte de gemelos 0,112 —	0,108 —	
7. Isacsson (1997)	Registro Sueco gemelos (hombres y mujeres). Gemelos mismo sexo nacidos 1926-1958. Datos administrativos ingresos (promedio de 3 años). Variable <i>background</i> familiar es educación gemelo	Gemelos idénticos 0,046 (0,001)	0,040 (0,001)	0,055 (0,002)
		Mellizos 0,047 (0,001)	0,046 (0,001)	0,054 (0,002)

Nota: <sup>a</sup> Véase texto de las fuentes y la información sobre los estudios individuales. <sup>b</sup> En este estudio la especificación VI trata a la educación y la experiencia como una variable endógena y utiliza las variables de *background* familiar y la edad como instrumentos. <sup>c</sup> En estas especificaciones la educación de cada gemelo se instrumenta con el reporte de la educación del otro gemelo.

El quinto, sexto y séptimo estudio presentado en la tabla 1.5 utilizan muestras de gemelos: en cada caso el *background* familiar es medido por la educación de un hermano. Curiosamente, el efecto de añadir la educación de un gemelo en estas muestras es similar al efecto de añadir el *background* de los padres en los otros estudios: el coeficiente de la educación cae en un 10-25%. Como los niveles de educación de los gemelos está incluso más altamente correlacionado que los niveles educativos entre padres-hijos o entre hermanos, la magnitud de esta caída no está lejos de la disminución atribuida sólo a los errores de medida. Los artículos de Miller *et al.* y Ashenfelter-Rouse permiten contrastar directamente la hipótesis del “error de medida puro”, ya que en ambos casos los autores presentan estimaciones para modelos por VI que incluyen los niveles educativos de los gemelos (a partir de los datos ofrecidos por uno de los gemelos) instrumentados por el nivel educativo del otro gemelo<sup>63</sup>. Como se muestra en la tabla, las estimaciones corregidas por error de medida del rendimiento de la educación con controles de la educación del gemelo son aproximadamente iguales a las correspondientes estimaciones MCO corregidas por errores de medida que no controlan *background* familiar.

Basándome en estos hallazgos para gemelos, y en los resultados de los otros estudios de la tabla 1.5, es posible concluir que cualquier sesgo que exista en las estimaciones convencionales del rendimiento de la educación por MCO, está también presente en modelos que controlan el *background* familiar. Aparte de un efecto atribuible al error de medida, el rendimiento de la educación es aproximadamente el mismo cuando se introducen controles de la educación de los padres o los hermanos. En el contexto de los modelos presentados en la tabla 1.3, este resultado sugiere que el componente de sesgo en el estimador MCO básico,  $\lambda_0 + \psi_0 \bar{S}$ , es aproximadamente del mismo tamaño, o solo ligeramente mayor, que el sesgo en el estimador que controla *background* familiar,  $\lambda_1 + \psi_1 \bar{S}$ .

Por otra parte, las medidas de *background* familiar tales como la educación de los padres o de los hermanos, generalmente ejercen un pequeño efecto positivo sobre los ingresos (es decir, el término  $\lambda_2 + \psi_2 \bar{S}$  es positivo). Por tanto, las estimaciones por VI que emplean el *background* familiar como un determinante exógeno de la educación esta a menudo (pero no siempre) considerablemente por encima de las correspondientes estimaciones MCO<sup>64</sup>. Esta conclusión es potencialmente importante para interpretar otras estima-

<sup>63</sup> De hecho Ashenfelter y Rouse (1998) incluyen estimaciones de modelos que incluyen  $S_{i1}$  y  $(S_{i1} + S_{i2})/2$  (es decir, la educación promedio de la familia). Estos coeficientes pueden ser “inconfundibles” para indicar el efecto directo de  $S_{i1}$  y  $S_{i2}$ , sin embargo, no hay suficiente información para construir los errores estándar para estos efectos.

<sup>64</sup> Esta conclusión difiere ligeramente de la conclusión provisional de Griliches (1979, p. 559) de que “las características de los padres... parecen afectar los ingresos principalmente a través de su efecto sobre el nivel de educación alcanzado.” Otra dimensión del *background* familiar que parece tener algún efecto sobre la educación de las mujeres es la composición de género de sus hermanos. Butcher y Case (1994) muestran que las mujeres con hermanos (en lugar de hermanas) tienen un poco más de educación. También utilizan la composición de género como instrumento de la educación y encuentran mayores estimaciones del rendimiento de la educación por VI que por MCO. A pesar de que la composición de género es aleatoria, no está claro que su único efecto sobre los ingresos sea a través de la educación: por lo tanto, las estimaciones VI de Butcher y Case (1994) pueden estar sesgadas.

ciones por VI del rendimiento de la educación basadas en factores como la proximidad a la universidad u otros aspectos institucionales del sistema educativo. En la medida en que los individuos de los grupos tratamiento y control de un análisis cuasi-experimental, tengan diferente *background* familiar, se podría esperar un sesgo al alza positivo en los estimadores VI obtenidos. Los resultados VI en la tabla 1.5 sugieren que es especialmente importante controlar el *background* familiar (o verificar que el *background* familiar es el mismo para los grupos tratamiento y control) en cualquier análisis de variables instrumentales del rendimiento de la educación.

Como se planteó en la sección 3, aunque la adición de controles de *background* familiar no necesariamente llevará a estimaciones consistentes de los verdaderos rendimientos educativos, bajo ciertos supuestos la estimación de modelos que controlan por *background* familiar pueden ser empleados para obtener estimadores consistentes del rendimiento marginal medio de la educación. Específicamente, si uno supone que  $\lambda_1 + \psi_1 \bar{S} = \lambda_2 + \psi_2 \bar{S}$ , el sesgo al alza en el coeficiente estimado de educación es igual al límite de probabilidad del coeficiente de la variable de *background* familiar. Por lo tanto, se puede restar el último del primero y obtener una estimación consistente de  $\bar{\beta}$ . Dado que las variables de *background* familiar como la educación de los padres o de los hermanos generalmente ejercen un pequeño efecto positivo sobre los ingresos, la aplicación de este procedimiento a los estudios en la tabla 1.5 conduciría a estimaciones del rendimiento marginal medio de la educación que están algo por debajo de las estimaciones MCO. Restar del coeficiente de educación de los padres o hermanos el coeficiente de educación del individuo es equivalente a un estimador *intra-familia*. Ya que los supuestos requeridos para justificar este estimador son más atractivos en el caso de los gemelos, pospongo a la siguiente sección una discusión más detallada.

Bajo supuestos ligeramente más débiles pero con más información —especialmente, con información de los ingresos del miembro de la familia cuyos datos son empleados como control— aún podría ser posible estimar el rendimiento marginal promedio de la educación. Particularmente, bajo el supuesto de “efectos familiares puros” de que hermanos o padres comparten la misma habilidad, se puede obtener una estimación de  $\bar{\beta}$  de los coeficientes de una regresión aparentemente no relacionada (SURE) de los ingresos de cada miembro de la familia sobre su propia educación, y la educación de los otros miembros de la familia (véase ecuaciones (17a) y (17b)). Ashenfelter y Zimmerman presentan estimaciones de este procedimiento aplicadas a hermanos y pares padres-hijos, con o sin correcciones del sesgo de errores de medida. Sus métodos de estimación ignoran la heterogeneidad en los rendimientos de la educación. Esto no es un problema para su muestra de hermanos, quienes tienen aproximadamente la misma educación media, pero sí lo puede ser para su muestra padres-hijos, ya que los hijos tienen alrededor de cuatro años de educación más que sus padres<sup>65</sup>.

<sup>65</sup> Si existe heterogeneidad en los rendimientos de la educación, las ecuaciones (17a) y (17b) sugieren que los coeficientes de la regresión SURE dependen del nivel educativo medio de los diferentes miembros de la familia.

Tras tener en cuenta sesgos plausibles de error de medida, los hallazgos de Ashenfelter y Zimmerman para hermanos implican estimaciones de  $\bar{\beta}$  aproximadamente iguales a las correspondientes estimaciones MCO. Sus estimaciones para pares de padres-hijos son más sensibles a los supuestos sobre si el verdadero rendimiento de la educación es el mismo para padres e hijos, y si los padres e hijos son intercambiables en las proyecciones lineales para el término de habilidad familiar latente. Sus especificaciones menos restrictivas sugieren una estimación de  $\bar{\beta}$  ligeramente inferior para los padres que las correspondientes estimaciones MCO, pero mucho menores estimaciones de  $\bar{\beta}$  para hijos. Sin embargo, dados sus resultados para hermanos, una interpretación alternativa es que el supuesto de “efectos familiares puros” es inapropiado para pares de padres-hijos. De hecho, Ashenfelter y Zimmerman encuentran que un modelo ligeramente modificado que permite que la habilidad familiar latente tenga un efecto diferencial sobre el intercepto de la ecuación de ingresos de padres e hijos, parece ajustar los datos bastante bien. Después de corregir por errores de medición, esta especificación sugiere que las estimaciones de  $\bar{\beta}$  para padres e hijos son un 25-50% menores que las correspondientes estimaciones por MCO.

### ■ 1.4.3. Estudios de educación e ingresos con gemelos

La tabla 1.6 resume 5 estudios recientes que comparan la educación y los ingresos de gemelos. Dos aspectos de estos estudios los diferencian de la literatura previa sobre gemelos revisada por Griliches (1979). Primero, las muestras en la literatura reciente son relativamente grandes, y tienden a incluir un rango más amplio de grupos de edad y de *background* familiar. Segundo, siguiendo el innovador artículo de Ashenfelter y Krueger (1994), la mayoría de estudios recientes abordan de lleno el problema del error de medida. Para cada estudio muestro un rendimiento de la educación de corte transversal (MCO), y dos estimaciones diferenciadas *intra-familia*: una estimada por MCO y otra corregida por errores de medida.

El estudio de Ashenfelter y Rouse (1998) emplea 3 años de datos del Princeton Twins Survey (PTS): su muestra incluye 340 parejas de gemelos idénticos, 60% de los cuales son mujeres. Como se muestra para las dos especificaciones de la tabla 1.6, las estimaciones del rendimiento de la educación *intra-familia* de Ashenfelter y Rouse, son aproximadamente un 30% menores que sus correspondientes estimaciones MCO. Este resultado contrasta con los resultados de Ashenfelter y Krueger (1994) basados en sólo 1 año de datos del PTS, que señalaban a un mayor efecto *intra-familia* que la estimación MCO<sup>66</sup>. El cuestionario de la PTS preguntaba a cada gemelo su nivel educativo, y el de su hermano. Este conjunto de respuestas adicionales permite a Ashenfelter y Rouse emplear la respuesta de uno de los gemelos sobre la diferencia en educación de su par como instrumento para las respuestas

<sup>66</sup> Rouse (1997, tabla 3) presenta algunos resultados que sugieren que los resultados de Ashenfelter y Krueger son atribuibles a la variabilidad muestral asociada a su relativamente pequeña muestra.

Tabla 1.6

### ESTIMACIONES DE CORTE TRANSVERSAL E INTRA-FAMILIA DIFERENCIADAS DE LA RENTABILIDAD DE LA EDUCACIÓN PARA GEMELOS<sup>a</sup>

AUTOR	MUESTRA Y ESPECIFICACIÓN		MCO	DIFERENCIADO	
			CORTE TRANSVERSAL	MCO	VI
1. Ashenfelter y Rouse (1998)	1991-1993 Encuesta gemelos Princeton Gemelos idénticos, hombres y mujeres. Controles Básicos incluye cuadrático En edad, sexo y raza. Controles añadidos incluye experiencia, estado civil Y estado sindical.	Básico	0,110 (0,010)	0,070 (0,019)	0,088 (0,025)
		Básico + adición	0,113 (0,010)	0,078 (0,018)	0,100 (0,023)
2. Rouse (1997)	1991-1995 Encuesta gemelos Princeton. Gemelos idénticos, hombres y mujeres Controles Básicos como anterior.		0,105 (0,008)	0,075 (0,017)	0,110 (0,023)
3. Miller <i>et al.</i> (1995)	Registro Australiano gemelos. Gemelos idénticos y mellizos. Controles incluye cuadrático en Edad, sexo y estado civil. Ingreso imputado de ocupación	Gemelos Idénticos	0,064 (0,002)	0,025 (0,005)	0,048 (0,010)
		Mellizos	0,066 (0,002)	0,045 (0,005)	0,074 (0,008)
4. Behrman <i>et al.</i> (1994)	NAS-NRC hombres blancos gemelos nacidos 1917-1927, más hombres gemelos nacidos 1936-1955 Registro gemelos Minnesota Controles incluye cuadrático en edad <sup>b</sup>	Gemelos Idénticos	0,071 (0,002)	0,035 (0,005)	0,056 —
		Mellizos	0,073 (0,003)	0,057 (0,005)	0,071 —
5. Isacsson (1997)	Gemelos Suecos del mismo sexo con Medidas de educación administrativas y de encuesta. Controles incluye sexo, estado civil, cuadrático en edad, y Residencia en una ciudad grande <sup>c</sup>	Gemelos Idénticos	0,049 (0,002)	0,023 (0,004)	0,024 (0,008)
		Mellizos	0,051 (0,002)	0,040 (0,003)	0,054 (0,006)
6. Isacsson (1997)	Gemelos Suecos del mismo sexo. Controles como anterior	Gemelos Idénticos	0,046 (0,001)	0,022 (0,002)	0,027/0,060 (0,003) (0,007)
		Mellizos	0,047 (0,002)	0,039 (0,002)	0,044/0,060 (0,002) (0,003)

*Nota:* <sup>a</sup>Véase texto de las fuentes y la información sobre los estudios individuales. <sup>b</sup>Estos autores no reportan estimaciones VI. Sin embargo, señalan una estimación de la fiabilidad de la diferencia en la escolarización de los gemelos idénticos de 0,62 (Behrman *et al.*, 1994, tabla 4). La estimación VI para los gemelos idénticos es la estimación diferenciada dividida por 0,62. No reportan una estimación de la fiabilidad para los mellizos. La estimación VI para los mellizos es la estimación diferenciada dividida por 0,80. <sup>c</sup>La diferencia en las estimaciones basadas en la educación registrada, es instrumentado por la diferencia en las medidas basadas en encuestas. <sup>d</sup>Las estimaciones VI reportados en esta fila son construidos por el supuesto de que la fiabilidad de la escolarización es 0,095 (aportando estimaciones bajas de la rentabilidad de la educación intra-familia) o 0,085 (aportando las mayores estimaciones).

del otro gemelo<sup>67</sup>. Las estimaciones VI, presentadas en la tercera columna de la tabla 1.6, son un 25% mayores que las estimaciones básicas diferenciadas, y alrededor de un 10% menor para la correspondiente estimación MCO. Rouse (1997) amplía el análisis hecho en Ashenfelter y Rouse, con una año adicional de datos del PTS. Sus resultados, resumidos en la fila 2 de la tabla 1.6, son generalmente consistentes con los de Ashenfelter y Rouse (1998), aunque las estimación VI de Rouse está algo por encima de la estimación presentada por Ashenfelter y Rouse, y de hecho supera la estimación MCO para la misma muestra<sup>68</sup>.

El estudio de Miller *et al.* (1995) emplea los datos de 1170 parejas de gemelos australianos (aproximadamente la mitad mujeres). La ventaja del gran tamaño muestral es compensar la ausencia de información de ingresos: Miller *et al.* tienen que imputar ingresos basados en la ocupación a dos dígitos. Así, gemelos con la misma ocupación a dos dígitos son codificados como si tuvieran el mismo ingreso<sup>69</sup>. Para gemelos idénticos, Miller *et al.* (1995) encuentran que la estimación *intra-familia* del rendimiento de la educación es casi un 50% menor que la estimación de corte transversal; para mellizos, el estimador *intra-familia* es un 40% menor. Como en la PTS, la base de datos de gemelos de Australia incluye múltiples informaciones de la educación de cada gemelo. Miller *et al.* (1995) sigue el procedimiento de Ashenfelter y Krueger (1994) de emplear las respuestas de uno de los gemelos sobre las diferencias en educación del par como un instrumento para las respuestas del otro gemelo. Para gemelos idénticos, las estimaciones VI resultantes están aproximadamente un 40% por encima de las estimaciones MCO diferenciadas, pero aún, un 25% por debajo de la estimación de corte transversal. Para mellizos la estimación por VI es de hecho ligeramente superior a la estimación MCO.

Behrman *et al.* (1994) analizan una base de datos que une la muestra de hombres blancos veteranos de la Segunda Guerra Mundial de la NAS-NRC, con datos de hombres del Minnesota Twins Registry<sup>70</sup>. Aunque el objeto principal de su estudio son los modelos de asignación inter-familiar de recursos, un cuadro del anexo presenta estimaciones de corte transversal y *intra-familia* del rendimiento educativo. Para gemelos idénticos Behrman *et al.*

<sup>67</sup> Con dos medidas de la educación de cada gemelo hay cuatro estimaciones posibles de las diferencias en educación. Ashenfelter y Krueger (1994) y Rouse (1997) examinan la estructura de covarianza de estas diferencias y concluyen que los errores de medida en la información de un gemelo dado de su propia educación y la de su hermano están ligeramente correlacionados. Las diferencias en las informaciones que un gemelo dado proporciona de los dos niveles educativos eliminará esta correlación.

<sup>68</sup> La estimación por VI de Rouse (1997) en la tabla 1.6 (que emplea la información aportada por uno de los gemelos sobre la diferencia en la educación de su par, como un instrumento para el otro gemelo) no se incluye en su estudio, pero Rouse se la comunicó en una conversación privada a Gary Solon.

<sup>69</sup> Sería interesante comparar el uso de los datos de ingresos reales y de los ingresos imputados en una base de datos que incluya ambos, como la PTS, para juzgar si la imputación afecta diferencialmente las estimaciones de la rentabilidad de la educación de corte transversal frente a la *intra-familia*.

<sup>70</sup> La muestra NAS-NRC ha sido ampliamente analizada por algunos de los mismos coautores, por ejemplo, Behrman *et al.* (1980). Behrman *et al.* imputan ingresos para la muestra de Minnesota a partir de la ocupación.

(1994) encuentra que la estimación *intra-familia* de la rentabilidad de la educación es cerca de un 50% mayor para las estimaciones MCO de corte transversal<sup>71</sup>, mientras para mellizos la proporción relativa es del 80%. Aunque estos autores no estiman realmente modelos por VI para corregir el error de medida, Behrman *et al.* (1994) muestran que la fiabilidad de la diferencia *intra-familia* en la educación para gemelos idénticos en la muestra de la NAS-NRC es 0,62. Con base en este cálculo, una estimación de la rentabilidad de la educación corregida por el efecto *intra-familia* para gemelos idénticos es de 0,056. Behrman *et al.* (1994) no ofrecen estimaciones comparables de la fiabilidad de esta relación para mellizos. Los resultados de Miller *et al.* (1995) y Ashenfelter y Krueger (1994), sin embargo, sugieren que la fiabilidad de las diferencias dentro de la familia en la educación de los mellizos es aproximadamente de 0,8. Según este cálculo, una estimación corregida de la rentabilidad de la educación para mellizos es de 0,071. Las magnitudes relativas de los estimadores MCO y *intra-familia* para gemelos y mellizos en Behrman *et al.* (1994) y Miller *et al.* (1995) son, por lo tanto, muy similares.

Finalmente, Isacsson (1997) analiza las diferencias de ingreso y educación entre una muestra grande de gemelos suecos (cerca de la mitad mujeres). Para una submuestra de datos dispone de información de dos medidas de educación: una del registro tomado por Estadística de Suecia, y otra basada en la auto-clasificación de nivel educativo<sup>72</sup>. Como se muestra en la tabla 1.6, Isacsson encuentra que el estimador *intra-familia* de la rentabilidad de la educación para gemelos idénticos en la submuestra con dos medidas de educación es menos del 50% de la magnitud del correspondiente estimador MCO, mientras para mellizos la proporción es del 80%. El autor construye estimaciones VI para el modelo *intra-familia* empleando la diferencia en las medidas de educación de la encuesta como instrumento de las diferencias en las medidas del registro<sup>73</sup>. Para gemelos idénticos, el estimador VI *intra-familia* está solo marginalmente por encima de la estimación MCO *intra-familia*, lo que implica la práctica inexistencia de sesgo de error de medida. Para mellizos, por otra parte, el procedimiento VI aumenta la estimación *intra-familia* en un 35%. Dado que se habría esperado una mayor atenuación del error de medida para gemelos que para mellizos, los patrones de los hallazgos de Isacsson son un poco paradójicos.

Isacsson (1997) también realiza estimaciones corregidas por errores de medida del rendimiento de la educación para una muestra más amplia de gemelos, suponiendo estimaciones “bajas” y “altas” de la fiabilidad de su principal medida de educación (confiabili-

<sup>71</sup> Esta proporción es ligeramente superior a la reportada en los primeros trabajos de Behrman *et al.* (1980) para gemelos idénticos con la muestra NAS-NRC. Griliches (1979) caracterizó sus resultados presentando una reducción del 65% en el rendimiento de la educación entre los estimadores MCO e *intra-familia*.

<sup>72</sup> Hay una diferencia fundamental de tiempo entre las dos medidas. Las estimaciones basadas en registros pertenecen a 1990, mientras que las medidas de auto-clasificación fueron recopilados en 1974. Los datos de ingresos de Isacsson se basan en los registros administrativos para los años 1987, 1990 y 1993.

<sup>73</sup> También presenta evidencia sobre la idoneidad de los supuestos que son necesarios para justificar la consistencia de estas estimaciones.

dades de 0,85 y 0,95, respectivamente). Los resultados son resumidos en las dos últimas filas de la tabla 1.6. Para mellizos las estimaciones *intra-familia* corregidas se sitúan en un rango bastante ajustado (0,044-0,060) que contiene el estimador VI *intra-familia* basado en dos medidas de educación (0,054). Para gemelos idénticos el rango de las estimaciones corregidas es más amplio (0,027-0,060) y se sitúa por encima de la estimación por VI *intra-familia* basada en las dos medidas de educación (0,024).

Vistos en conjunto, los resultados de Isacson sugieren que la estimación del rendimiento de la educación *intra-familia* corregida por errores de medida, para mellizos en Suecia es aproximadamente tan grande o incluso mayor que la correspondiente estimación MCO. La precisión de la magnitud de la estimación *intra-familia* corregida de errores de medida para gemelos idénticos es más incierta, y parece muy sensible a los supuestos sobre el error de medida. Una interpretación prudente de los resultados de Isacson es que puede haber un sesgo al alza en las estimaciones MCO del rendimiento de la educación respecto a la estimación *intra-familia* para gemelos idénticos.

¿Qué conclusión general se puede extraer de la reciente literatura sobre gemelos? Supongamos que a priori creemos que gemelos idénticos tienen habilidades idénticas. Entonces el estimador *intra-familia* para gemelos idénticos, corregido por sesgo de error de medida, es consistente para el rendimiento marginal promedio de la educación para la toda la población de gemelos<sup>74</sup>. Suponiendo que este es el caso, las estimaciones de la tabla 1.6 sugieren que el estimador MCO de corte transversal aporta una estimación ligeramente sesgada al alza de los rendimientos marginales promedio de la educación: la magnitud del sesgo en los estudios oscila desde el 50% (Isacson) a cero (Rouse, 1997). Dadas las limitaciones de los datos de ingresos imputados utilizados por Miller *et al.* (1995) y Behrman *et al.* (1994), y la incertidumbre en las correcciones del error de medida para el estudio de Isacson, concedo mayor validez a los estudios de Ashenfelter-Rouse y Rouse, los cuales sugieren un rango menor del sesgo —aproximadamente 10-15%.

Una segunda conclusión emerge de los tres estudios que presentan resultados para mellizos. En estos estudios el estimador *intra-familia* del rendimiento de la educación corregido de errores de medida para mellizos es aproximadamente igual al correspondiente estimador MCO. Curiosamente, la estimación del rendimiento de la educación de Ashenfelter y Zimmerman, corregida de errores de medida para hermanos —construida bajo el supuesto de que los hermanos tienen habilidades idénticas— es también, aproximadamente igual a la correspondiente estimación MCO. Dado que los mellizos son básicamente hermanos (o hermanas) con la misma edad, la similitud de estos hallazgos para mellizos y hermanos contribuye a su validación. Suponiendo que las estimaciones MCO están sesgadas al alza respecto al verdadero efecto causal promedio de la educación, las estimaciones *intra-familia* basadas en mellizos y hermanos deben estar también sesgada al alza. Además, dado que el estimador MCO está sesgado a la baja por el error de medida, mientras que las estimaciones *intra-familia* corregidas

<sup>74</sup> La “población de gemelos” puede ser razonablemente extensa o muy reducida, según la base de datos.

para mellizos o hermanos no lo están, se puede concluir que el sesgo de habilidad en los estimadores *intra-familia* para mellizos o hermanos es más pequeño que el sesgo de habilidad en los estimadores MCO de corte transversal: del orden de la mitad<sup>75</sup>. Esto implica que las diferencias en habilidad entre hermanos o hermanas son determinantes relativamente menos importantes del nivel de educación *intra-familia*, de lo que son las diferencias de habilidad total en la determinación del nivel de educación para la población en general.

Este hallazgo plantea la interesante cuestión de cómo y por qué las familias afectan las decisiones de escolarización de los niños con habilidades diferenciales. Behrman *et al.* (1982) presentan un modelo que incorpora las preferencias de los padres en la distribución de los recursos educativos entre hermanos que es consistente con el comportamiento reforzador o compensatorio (es decir, las familias pueden gastar más en la educación de sus hijos más hábiles o menos hábiles). Sus hallazgos empíricos sustentan la noción del comportamiento compensador de los padres —comportamiento que conduciría a una disminución en la importancia relativa de las diferencias en habilidades en la determinación del nivel de educación dentro de las familias más que entre las familias<sup>76</sup>.

Si no se cree que los gemelos idénticos tienen habilidades idénticas, entonces incluso el estimador *intra-familia* del rendimiento de la educación para gemelos idénticos puede estar sesgado por diferencias de habilidad. Ashenfelter y Rouse (1998) presentan una diversa evidencia indirecta que apoya la hipótesis de que los gemelos idénticos son realmente idénticos, y que las diferencias en sus niveles de educación son atribuibles a factores aleatorios y no a diferencias de habilidad. Por ejemplo, señalan que las diferencias de educación entre los gemelos idénticos no se correlacionan con el orden de nacimiento y con la educación de su cónyuge<sup>77</sup>. A pesar de esta evidencia, y el fuerte atractivo intuitivo del supuesto de “igualdad de habilidades” para los gemelos idénticos, sospecho, no obstante, que los investigadores con una fuerte creencia a priori en la importancia del sesgo de habilidad no serán convencidos.

#### ■ 1.4.4. Evidencia directa de la heterogeneidad en los rendimientos educativos

Un último conjunto de resultados en la literatura reciente que cabe reseñar brevemente se refiere a fuentes observables de variación del rendimiento de la educación. Entre

<sup>75</sup> Escribimos el plim del estimador MCO como  $R_0(\bar{\beta} + G_{mco})$  y el plim del estimador basado en mellizos corregido del error de medida como  $(\bar{\beta} + G_I)$ . Si los estimadores MCO y corregido para mellizos son aproximadamente iguales, entonces  $R_0(\bar{\beta} + G_{mco}) = (\bar{\beta} + G_I)$ . Suponiendo que  $R_0 \approx 0,9$ , si  $G_I$  es entre un 10-20% de  $\bar{\beta}$ ,  $G_I/G_{mco}$  está entre un 45-60%.

<sup>76</sup> El estudio más reciente de Behrman *et al.* (1994), no obstante, encuentra comportamiento reforzador en la asignación de recursos escolares dentro de las familias.

<sup>77</sup> Este último hallazgo parece estar en contradicción con los resultados de Behrman *et al.* (1994, tabla A2) que muestra una fuerte relación entre las diferencias en educación y las diferencias en la educación de los cónyuges entre los gemelos idénticos.

las fuentes potenciales de heterogeneidad que han sido identificadas y estudiadas están la calidad de la escuela, el *background* familiar, y la habilidad, medidas por el Cociente Intelectual (CI) o puntuaciones de pruebas de aptitud.

Gran parte del interés en la conexión entre calidad escolar y rentabilidad de la educación fue estimulado por la observación de que los hombres negros tenían un rendimiento de la educación considerablemente menor que los hombres blancos a principios de los sesenta (por ejemplo, Welch, 1973). Además, la mayor parte de la convergencia en los salarios relativos entre negros y blancos que tuvo lugar en los sesenta y setenta puede ser atribuida a la combinación de unos rendimientos de la educación relativos crecientes para las cohortes más recientes de hombres negros, y la creciente educación relativa de los negros respecto a los blancos (Smith y Welch, 1986; 1989). Como la calidad relativa de las escuelas a las que asistían los negros en los segregados estados del sur mejoraron significativamente entre 1920 y 1960 (Card y Krueger, 1992b), estos factores llevaron a los investigadores a especular que los incrementos en la calidad escolar puede generar incrementos en el nivel educativo y mayores rendimientos a la educación.

Card y Krueger (1992a,b) estiman tasas de rendimiento de la educación para diferentes cohortes de hombres blancos y negros que nacieron en diferentes estados y correlacionan estos rendimientos con medidas de calidad escolar por cohorte y estado de nacimiento<sup>78</sup>. Una característica diferenciadora de los rendimientos de la educación que complica este análisis es el hecho de que los diferenciales salariales asociados a la educación son mayores en algunas partes de EE.UU. que en otras<sup>79</sup>. Card y Krueger intentan resolver esta dificultad suponiendo una estructura aditiva del rendimiento educativo: un individuo nacido en un estado y trabajando en otro recibe la suma del componente estado de nacimiento (que presumiblemente varía con la calidad escolar); y el componente del estado de residencia<sup>80</sup>. Bajo estos supuestos, Card y Krueger (1992a,b) muestran que el componente del estado de nacimiento en el rendimiento de la educación está sistemáticamente correlacionado con las características del sistema escolar. Por ejemplo, sus resultados sugieren que la disminución en 10 estudiantes de la ratio alumnos-profesor por estado, incrementa la tasa de rentabilidad de la educación obtenida por los estudiantes del estado en cerca de 0,9 puntos porcentuales.

Desde el punto de vista de los modelos presentados en la sección 3, otro interesante hallazgo planteado por Card y Krueger es que los estudiantes que crecen en estados con

<sup>78</sup> Dentro de los estudios anteriores que modelan el efecto de la calidad escolar sobre el rendimiento de la educación están los de Akin y Garfinkel (1980) y Link *et al.* (1980).

<sup>79</sup> Esta característica del mercado laboral estadounidense es documentado por Chiswick (1974). Dahl (1997) presenta un resumen detallado de los cambios en la rentabilidad de la educación por estado en 1980 y 1990, y evalúa la contribución de la migración selectiva a estos patrones.

<sup>80</sup> Este supuesto es criticado por Heckman *et al.* (1996), porque ignora la posibilidad de migración selectiva. Curiosamente, Heckman *et al.* encuentran mayores efectos promedio de calidad escolar en los modelos que controlan por migración selectiva mediante la inclusión de una función de la distancia que las personas han emigrado entre el Estado de residencia y Estado de nacimiento. Véase Card y Krueger (1996) para un resumen y discusión.

mejor calidad escolar adquieren más educación. Por ejemplo, sus resultados para hombres blancos sugieren que una reducción de 10 estudiantes en la ratio alumnos-maestros por estado, aumenta el nivel educativo promedio en 0,6 años. En principio, la calidad escolar puede afectar el nivel educativo reduciendo el coste marginal de la educación, o aumentando los beneficios marginales de la educación, o por ambos. Si se ignora el efecto coste, entonces la estimación implícita del parámetro  $k$  en la ecuación (4) para hombres blancos nacidos en el periodo 1920-1950 es de 0,013<sup>81</sup>. Esto a su vez sugiere que la magnitud del componente de endogeneidad ( $\psi_0 \bar{S} = k\bar{f}\bar{S}$ ) en la estimación MCO del rendimiento de la educación es aproximadamente  $0,15f$ , donde  $f$  es la proporción de la varianza en los resultados educativos que es atribuible a las diferencias en habilidad, frente a la atribuible a las diferencias en preferencias. Suponiendo que el sesgo de endogeneidad es de un 0,015 (como lo sugieren los resultados de Ashenfelter y Rouse, 1998)  $f$  es aproximadamente del 10%. Estos cálculos son evidentemente especulativos: no obstante, ilustran la utilidad potencial de los datos sobre los determinantes observables del rendimiento de la educación en el desarrollo de una mejor comprensión del efecto causal de la educación.

Evidencia adicional sobre el grado de heterogeneidad en los rendimientos de la educación y su relación con la calidad escolar y el *background* familiar es presentado en una serie de artículos de Altonji y Dunn (1995, 1996a,b) que estudian los datos de ingresos y educación para parejas de hermanos en el *National Longitudinal Surveys of Young Men and Young Women* y el *Panel Study of Income Dynamics*. Altonji y Dunn ajustan modelos del logaritmo de los ingresos que incluyen educación, una serie de variables de control, e interacciones de la educación con la educación de los padres, cociente intelectual, y características de calidad escolar<sup>82</sup>. Los autores estiman estos modelos excluyendo e incluyendo efectos hijos de familia. Estas últimas especificaciones es quizá el aspecto más interesante de su trabajo, dado que en estos modelos los efectos directos o principales del *background* familiar se mantienen constantes. Como se esperaría de la discusión de los modelos de hermanos y gemelos de la sección 3, los errores de medida cumplen un papel importante en los modelos *intra-familia*: Altonji y Dunn presentan estimaciones de la magnitud probable de los sesgos de atenuación que aparecen en estos modelos, y consecuentemente interpretan sus estimaciones.

Los resultados de Altonji y Dunn sugieren que una mayor calidad escolar, medida por el gasto por alumno, salarios medios de los maestros o un índice compuesto, aumenta el rendimiento educativo. Respecto al *background* familiar y la habilidad sus resultados son

<sup>81</sup> En el modelo,  $S_i = (b_i - r_i)/k$ . Si un aumento de la calidad educativa que incrementa la rentabilidad promedio de la educación en un 0.009 conduce a un aumento de 0,6 en los años de educación, entonces  $k \approx 0,013$ .

<sup>82</sup> Hay una literatura preliminar, que incluye interacciones del *background* familiar y las medidas de habilidad con la educación. Hauser (1973) encontró escasa evidencia de que el nivel ocupacional del padre afecte el rendimiento de la educación de los hijos. Del mismo modo, Olneck (1979) concluye que el cociente intelectual y la educación del padre tienen poco efecto sistemático sobre la rentabilidad de la educación. Por otra parte, Hause (1972) y Willis y Rosen (1979) encuentran correspondencia positiva entre los resultados de las pruebas de aptitud y la educación.

menos concluyentes. En algunos de sus modelos donde incluyen efectos fijos de familia encuentran que la mayor educación de la madre aumenta el rendimiento educativo, aunque en otras muestras y especificaciones los efectos son débiles e incluso con el signo contrario. Al igual que la literatura inicial, encuentran efectos pequeños y no sistemáticos de la educación de los padres sobre los rendimientos de la educación en modelos que excluyen efectos fijos de familia. Los efectos del cociente intelectual sobre la rentabilidad de la educación son generalmente positivos (pero estimados de forma imprecisa) en los modelos *intra-familia* pero negativos en los modelos que excluyen efectos familiares.

Ashenfelter y Rouse (1998) también analizan los efectos del *background* familiar sobre la rentabilidad de la educación para gemelos idénticos. Al igual que Altonji y Dunn, sus estimaciones de las interacciones entre educación de los padres y la diferencia en la educación entre gemelos idénticos son positivas pero imprecisas.

Finalmente, Ashenfelter y Rouse (1998) presentan evidencia interesante sobre la existencia de rendimientos marginales de la educación decrecientes (es decir, concavidad en la relación entre el logaritmo de los ingresos y la educación a nivel individual). Los autores aumentan una sencilla ecuación de ingresos *intra-familia* para gemelos idénticos, con una interacción entre la educación promedio de los gemelos y sus diferencias en educación. En el contexto del modelo representado por la ecuación (15), el coeficiente sobre esta interacción es una estimación del coeficiente  $k_i^{83}$ . Ashenfelter y Rouse hallan que la rentabilidad de la educación disminuye con el nivel promedio de educación —desde cerca de 0,12 para 9 años de educación hasta 0,08 para 16 años de educación— aunque el gradiente no se mide de forma precisa. Esta evidencia directa de disminución de los rendimientos marginales de la educación apoya la interpretación de los estimadores VI en la tabla 1.4 como ofreciendo estimaciones del rendimiento marginal de la educación para individuos que en otro caso tendrían nivel de educación inferior al promedio (respecto a la población analizada en cada estudio).

Esta breve revisión sugiere tres conclusiones principales. Primero, el rendimiento de la educación está relacionado con algunas variables observables, como la raza, la calidad escolar, las medidas de *background* familiar, y quizás la habilidad medida. Segundo, aspectos como la raza, la calidad escolar, y la educación de la madre, que están asociados con mayor rendimiento de la educación, están también generalmente asociados con mayores niveles educativos. Estos patrones son compatibles con un modelo de optimización de la calidad escolar en el cual los individuos es más probable que elijan mayores niveles de educación si el rendimiento de la educación es mayor. Tercero, pero más tentativamente, los rendimientos individuales de la educación son decrecientes con el nivel de educación.

<sup>83</sup> Si se asume la igualdad en las habilidades para los gemelos idénticos, la ecuación (15) implica que  $\Delta \log y_i = b \Delta S_i - 1/2k_i (S_{i1}^2 - S_{i2}^2) \approx b \Delta S_i - k_i \bar{S}_i \Delta S_i$ , donde  $\bar{S}_i$  es la educación promedio de los gemelos en la familia  $i$ .

## ■ 1.5. CONCLUSIONES

Tomada en conjunto, considero que la literatura reciente sobre rendimientos de la educación apunta a cinco conclusiones fundamentales:

1. Consistente con el resumen de literatura de los sesenta y setenta elaborado por Griliches (1977, 1979), la rentabilidad media (o marginal medio) de la educación en una población dada no es muy inferior a la estimación resultante de una estimación sencilla de corte transversal de los ingresos sobre la educación. La “mejor evidencia disponible” de los últimos estudios de gemelos idénticos sugiere un pequeño sesgo al alza (del orden del 10%) en la estimación básica por MCO.
2. Las estimaciones de los rendimientos de la educación basados en comparaciones de hermanos o mellizos, contienen cierto sesgo de habilidad positivo, pero menos que la correspondiente estimación MCO. Las diferencias de habilidad parecen ejercer una influencia relativamente menor sobre las diferencias educativas dentro de la familia que sobre las diferencias entre familias.
3. Las estimaciones VI del rendimiento de la educación basados en el *background* familiar son sistemáticamente mayores que las correspondientes estimaciones MCO, y probablemente contienen un mayor sesgo de habilidad al alza que las estimaciones MCO.
4. Los rendimientos de la educación varían entre la población de acuerdo con factores observables tales como la calidad escolar y la educación de los padres.
5. Estimaciones por VI del rendimiento de la educación basadas en las intervenciones en el sistema educativo tienden a estar un 20% o más por encima de las correspondientes estimaciones MCO. Aunque se han aportado varias explicaciones de hallazgo, una hipótesis plausible es que rendimientos marginales de la educación de ciertos subgrupos de población —particularmente, aquellos subgrupos cuyas decisiones educativas están más afectadas por innovaciones estructurales en el sistema educativo— son algo mayores que los rendimientos marginales medios para la población en su conjunto.

Aunque la investigación en la última década ha hecho un verdadero progreso sobre la cuestión del efecto causal de la educación, puede ser útil concluir con una breve lista de temas relacionados que no han sido tan exhaustivamente abordados. Una cuestión no resuelta, es si el rendimiento privado a la educación —que es el foco del trabajo microeconómico presentado aquí— es igual, mayor o menor que los rendimientos sociales. Esta pregunta se sitúa en el centro de la controversia de la contabilidad del crecimiento que estimuló gran parte de la literatura moderna sobre el rendimiento de la educación, y ha vuelto a surgir en la década pasada con el regreso del interés sobre las fuentes de crecimiento económico de largo plazo. De hecho, gran parte de la “nueva” teoría del crecimiento, se cen-

tra en la posible existencia de externalidades a la educación significativas<sup>84</sup>. El estudio de externalidades a nivel agregado, es evidentemente más difícil que el estudio de la rentabilidad privada de la educación a nivel individual: no existen “gemelos idénticos” a nivel agregado. No obstante, algunas de las ideas que subyacen a los estudios cuasi-experimentales sobre la rentabilidad privada de la educación, pueden ser útiles a un nivel más agregado. Por ejemplo, cambios institucionales en el sistema educativo pueden generar ajustes en la oferta relativa de los trabajadores mejor educados en una región, respecto a otra que puede ser usado para construir contrastes cuasi-experimentales a nivel de mercado.

Una segunda cuestión (y relacionada), es si la rentabilidad privada de la educación opera a través de ajustes homogéneos en la productividad de los trabajadores mejor educados, o a través de mecanismos más complejos, como el acceso diferencial a distintos tipos de empleos. Algunos autores interpretan los estudios sobre el efecto credencialismo descritos en la sección 2 como la distinción entre estas alternativas (véase, por ejemplo, Weiss, 1995). Un innovador estudio de Tyler *et al.* (1998) sobre la rentabilidad del mercado de un certificado de Desarrollo Educacional General (DEG), sugiere que las credenciales tienen en sí mismas un valor significativo en el mercado laboral de EE.UU., mientras otros trabajos (por ejemplo, el de Cameron y Heckman, 1993) han cuestionado esta hipótesis<sup>85</sup>.

Un tercer aspecto, que ha recibido un renovado interés en la literatura reciente es como el rendimiento de la educación cambia con las características observables, como el *background* familiar, la calidad escolar, la habilidad o la localización. Uno de los objetivos dignos de investigación futura, es el desarrollo de una mejor comprensión de la medida en que el efecto de características permanentes como el *background* familiar sobre el rendimiento de la educación “explican” sus efectos sobre el nivel educativo. Un objetivo más elevado, es entender la determinación conjunta del nivel educativo y otros resultados endógenos, como la localización o la ocupación en el contexto de un modelo estructural de la determinación de la educación y los ingresos.

Un tema final que no he considerado es este capítulo es el cambio en el rendimiento de la educación en el tiempo: para la economía en su conjunto o para cohortes fijas de individuos. En los últimos 15 años, la medida convencional de rentabilidad de la educación ha crecido cerca de un 35-50% (véase, Autor *et al.*, 1997 o el capítulo de Katz y Autor en este volumen). Respecto a estos cambios, los sesgos de habilidad que son el foco de la literatu-

<sup>84</sup> Véase el capítulo de Topel en este volumen para un resumen de esta literatura con un énfasis en cuestiones relativas al capital humano.

<sup>85</sup> El diseño de la investigación de Tyler *et al.* (1998) pone de relieve el valor del conocimiento institucional detallado para ayudar a desentrañar los mecanismos causales en el mercado laboral. El certificado GED se otorga en lugar de la titulación de educación secundaria, por la exitosa culminación de un examen. En algunos estados, sin embargo, la puntuación requerida para obtener el GED es menor, permitiendo comprobar si el certificado, en sí mismo, es recompensado en el mercado de trabajo, o sólo el “conocimiento” implícito. Los resultados de Tyler *et al.* sugieren que el certificado en sí mismo es importante, dado que personas con la misma puntuación que obtendrían el título en un Estado pero en otro no, parecen ganar más cuando tienen el grado.

ra presentada aquí, parece bastante modesto en magnitud. No obstante, algunos autores han argumentado que los cambios en el tiempo en la rentabilidad de la educación, pueden ser generados en parte por *cambios* en la magnitud de los componentes del sesgo de habilidad (por ejemplo, Taber, 1998; Cawley *et al.*, 1998). Algunos de los métodos desarrollados para estudiar la importancia del sesgo de habilidad en una base de datos de corte transversal pueden ser extendidos a datos de panel, ofreciendo la posibilidad de modelar sesgos de habilidad cambiantes en el tiempo.

## ■ APÉNDICE A

### ■ A.1. Estimación MCO

Sea  $y$  una variable observable (escalar), que está relacionada con la covariable  $X$  de dimensión  $k$ , a través de un modelo de regresión lineal con intercepto aleatorio  $\alpha$  y coeficiente aleatorio de pendiente  $\beta$ :

$$y = \alpha + X'\beta + u = \bar{\alpha} + X'\bar{\beta} + (\alpha - \bar{\alpha}) + X'(\beta - \bar{\beta}) + u \quad (\text{A.1})$$

donde  $\bar{\alpha}$  y  $\bar{\beta}$  representan las medias de  $\alpha$  y  $\beta$ , respectivamente, y  $E[X'u] = 0$ . Denotamos la proyección lineal de  $\alpha$  y  $\beta$  sobre  $X$  como:

$$\alpha - \bar{\alpha} = \lambda'(X - \bar{X}) + v_1, \quad (\text{A.2a})$$

$$\beta - \bar{\beta} = \psi(X - \bar{X}) + v_2, \quad (\text{A.2b})$$

donde  $E[X'v_1] = E[X'v_2] = 0$  (por definición de  $\lambda$  y  $\psi$ ). Usando estas definiciones,

$$E[(\beta - \bar{\beta})(X - \bar{X})] = \psi E[(X - \bar{X})(X - \bar{X})'] = \psi \text{var}[X],$$

y por lo tanto,

$$\begin{aligned} \text{cov}[X, X'(\beta - \bar{\beta})] &= E[(X - \bar{X})(\beta - \bar{\beta})X'] = E[(X - \bar{X})(\beta - \bar{\beta})'(X - \bar{X})] \\ &= \text{var}[X]\psi\bar{X}' - D, \end{aligned}$$

donde

$$D = E[(X - \bar{X})(X - \bar{X})'(\beta - \bar{\beta})].$$

El límite de probabilidad del estimador MCO de  $\bar{\beta}$  para la ecuación (A.1) es por tanto,

$$\begin{aligned} \text{var}[X]^{-1} \text{cov}[X, y] &= \text{var}[X]^{-1} \{ \text{var}[X]\beta + \text{var}[X]\lambda + \text{var}[X]\psi'\bar{X} + D \} \\ &= \bar{\beta} + \lambda + \psi'\bar{X} + \text{var}[X]^{-1}D \end{aligned}$$

Nótese que si  $X$  y  $\beta$  están distribuidos simétricamente de forma conjunta, entonces  $D = 0$ . En este caso, el límite de probabilidad del coeficiente de regresión MCO es solo:

$$\text{plim}(\beta_{\text{ols}}) = \bar{\beta} + \lambda + \psi'\bar{X} \quad (\text{A.3})$$

## ■ A.2. Estimación del modelo de coeficientes aleatorios

Considere la estimación de la ecuación (A.1) cuando el conjunto de instrumentos  $Z$  está disponible con las siguientes propiedades:

$$E[(\alpha - \bar{\alpha}) | Z] = 0, \quad (\text{A.3a})$$

$$E[(\beta - \bar{\beta}) | Z] = 0. \quad (\text{A.3b})$$

Suponiendo que  $Z$  incluye un vector de constantes, denotemos la forma reducida de la proyección de  $X$  respecto a  $Z$  como:

$$X = \Pi Z + \nu \quad (\text{A.4})$$

Finalmente, supongamos que

$$E[u | X, Z] = 0, \quad E[\nu | Z] = 0 \quad (\text{A.5a})$$

$$E[\alpha - \bar{\alpha} | X, Z] = \lambda'_x X + \lambda'_z Z \quad (\text{A.5b})$$

$$E[\beta - \bar{\beta} | X, Z] = \psi_x X + \psi_z Z \quad (\text{A.5c})$$

El supuesto (A.5a) fortalece la condición de ortogonalidad, definiendo los componentes del error  $u$  y  $\nu$ , en los supuestos sobre las esperanzas condicionales. Los supuestos (A.5b) y (A.5c) especifican que las esperanzas condicionales de  $\alpha$  y  $\beta$  son lineales en  $X$  y  $Z$ . Bajo estos supuestos,

$$0 = E[\alpha - \bar{\alpha} | Z] = E[\lambda'_x (\Pi Z + \nu) + \lambda'_z Z | Z] = (\lambda'_x \Pi + \lambda'_z) Z,$$

lo que implica que  $\lambda'_z = -\lambda'_x \Pi$ . De la misma manera,

$$0 = E[\beta - \bar{\beta} | Z] = E[\psi_x (\Pi Z + \nu) + \psi_z Z | Z] = (\psi_x \Pi + \psi_z) Z,$$

implicando que  $\psi'_z = -\psi'_x \Pi$ . Al sustituir (A.5b) y (A.5c) en (A.1), y tomando esperanzas condicionales sobre  $(X, Z)$ , obtenemos:

$$\begin{aligned} E[y | X, Z] &= \bar{\alpha} + X' \bar{\beta} + \lambda'_x X + \lambda'_z Z + X' (\psi_x X + \psi_z Z) \\ &= \bar{\alpha} + X' \bar{\beta} + \lambda_x (X - \Pi Z) + X' \psi_x (X - \Pi Z) \\ &= \bar{\alpha} + X' \bar{\beta} + \lambda_x \nu + X' \psi_x \nu. \end{aligned} \quad (\text{A.6})$$

Aplicando los argumentos estándar, (A.6) supone que la estimación consistente de  $\bar{\beta}$  puede ser obtenida de un estimador "función de control" que incluye  $X$ ,  $\bar{\nu}$  (el residuo de la regresión de  $X$  sobre  $Z$ ), y las interacciones de  $X$  y  $\bar{\nu}$  (véase además Garen, 1984). Obsérvese que si es  $\beta$  constante, entonces la función de control es simplemente  $\bar{\nu}$ , aportando el tradicional estimador por VI. (En este caso, los supuestos anteriores pueden ser debilitados al reemplazar el operador esperanzas en la ecuación (A.3) y (A.5), por el operador de proyección lineal).

### ■ A.3. Errores de Medida en un modelo de regresión bivalente

Consideremos el modelo de regresión bivalente:

$$y = X_1 b_1 + X_2 b_2 + u, \quad (\text{A.7})$$

donde  $X_1$  y  $X_2$  se miden con error. Denotemos el valor observado de  $X_i$  por  $X_i^o$  ( $i = 1, 2$ ), y supongamos que:

$$X_1^o = X_1 + \varepsilon_1, \quad X_2^o = X_2 + \varepsilon_2,$$

donde  $E[X_i \varepsilon_j] = E[\varepsilon_1 \varepsilon_2] = 0$ . Dado que  $R_1$  y  $R_2$  representan la razón de fiabilidad de  $X_1$  y  $X_2$  respectivamente, donde:

$$R_i = \text{cov}[X_i^o, X_i] / \text{var}[X_i^o]$$

Finalmente, consideremos las regresiones auxiliares:

$$X_1 = X_1^o a_{11} + X_2^o a_{12} + v_1, \quad (\text{A.8a})$$

$$X_2 = X_1^o a_{21} + X_2^o a_{22} + v_2, \quad (\text{A.8a})$$

donde  $v_i$  es ortogonal a  $X_1^o$  y  $X_2^o$  para  $i = 1, 2$ . Los coeficientes de estas regresiones pueden ser expresados en términos de las varianzas de las  $X$  observadas, la razón de fiabilidad, y  $\rho$ , la correlación de las variables observadas,  $X_1^o$  y  $X_2^o$ . Si  $y$  se expresa como la regresión sobre las  $X$  observadas:

$$y = X_1^o c_1 + X_2^o c_2 + e,$$

los coeficientes de la regresión serán iguales a:

$$c_1 = b_1 a_{11} + b_2 a_{21}, \quad c_2 = b_1 a_{12} + b_2 a_{22},$$

Es fácil demostrar que

$$c_1 = b_1 \frac{R_1 - \rho^2}{1 - \rho^2} + b_2 \frac{1 - R_2}{1 - \rho^2} \times \frac{\text{cov}[X_1^o, X_2^o]}{\text{var}[X_1^o]}, \quad (\text{A.9a})$$

$$c_2 = b_2 \frac{R_2 - \rho^2}{1 - \rho^2} + b_1 \frac{1 - R_1}{1 - \rho^2} \times \frac{\text{cov}[X_1^o, X_2^o]}{\text{var}[X_2^o]}, \quad (\text{A.9b})$$

## ■ BIBLIOGRAFÍA

- AKIN, JOHN S. e IRWIN GARFINKEL (1980), "The quality of education and cohort variation in black-white earnings differentials: comment", *American Economic Review* 70: 186-191.
- ALTONJI, JOSEPH G. y THOMAS A. DUNN (1995), "The effects of school and family characteristics on the return to education", Working paper no. 5072 (NBER, Cambridge, MA).
- (1996a), "Using siblings to estimate the effect of school quality on wages", *Review of Economics and Statistics* 78: 665-671.
- (1996b), "The effects of family characteristics on the return to schooling", *Review of Economics and Statistics* 78: 692-704.
- ANGRIST, JOSHUA D. y GUIDO W. IMBENS (1995), "Two-stage least squares estimation of average causal effects in models with variable treatment intensity", *Journal of the American Statistical Association* 90: 431-442.
- ANGRIST, JOSHUA D. y ALAN B. KRUEGER (1991), "Does compulsory school attendance affect schooling and earnings?", *Quarterly Journal of Economics* 106: 979-1014.
- (1992), "Estimating the payoff to schooling using the Vietnam-era draft lottery", Working paper no. 4067 (NBER, Cambridge, MA).
- (1995), "Split sample instrumental variables estimates of the return to schooling", *Journal of Business and Economic Statistics* 13: 225-235.
- ANGRIST, JOSHUA D. y VICTOR LAVY (1997), "Using Maimonides' rule to estimate the effect of class size on scholastic achievement", Working paper no. 5888 (NBER, Cambridge, MA).
- ANGRIST, JOSHUA D. y WHITNEY K. NEWEE (1991), "Over-identification tests in earnings functions with fixed effects", *Journal of Business and Economic Statistics* 9: 317-323.
- ANGRIST, JOSHUA D.; GUIDO W. IMBENS y DONALD B. RUBIN (1996), "Identification of causal effects using instrumental variables", *Journal of the American Statistical Association* 91: 444M72.
- ASHENFELTER, ORLEY y COLIN HARMON (1998), "Editors introduction", *Labour Economics (special issue on education)*, in press.
- ASHENFELTER, ORLEY y ALAN B. KRUEGER (1994), "Estimates of the economic return to schooling for a new sample of twins", *American Economic Review* 84:1157-1173.
- ASHENFELTER, ORLEY y CECILIA E. ROUSE (1998), "Income, schooling and ability: evidence from a new sample of identical twins", *Quarterly Journal of Economics* 113: 253-284.
- ASHENFELTER, ORLEY y DAVID ZIMMERMAN (1997), "Estimates of the return to schooling from sibling data: fathers, sons and brothers", *Review of Economics and Statistics* 79: 1-9.
- AUTOR, DAVID H.; LAWRENCE F. KATZ y ALAN B. KRUEGER (1997), "Computing inequality: have computers changed the labor market?" Working paper no. 5956 (NBER, Cambridge, MA).
- BECKER, GARY S. (1964), *Human capital: a theoretical and empirical analysis, with special reference to education* (Columbia University Press, New York).
- (1967), *Human capital and the personal distribution of income* (University of Michigan Press, Ann Arbor, MI).
- BEHRMAN, JERE R.; Z. HRUBEC; PAUL TAUBMAN y TERENCE J. WALES (1980), *Socioeconomic success: a study of the effects of genetic endowments, family environment and schooling* (North Holland, Amsterdam).

- BEHRMAN, JERE R.; ROBERT POLLACK y PAUL TAUBMAN (1982), "Parental preferences and the provision of progeny", *Journal of Political Economy* 90: 52-73.
- BEHRMAN, JERE R.; MARK R. ROSENZWEIG y PAUL TANBMAN (1994), "Endowments and the allocation of schooling in the family and the marriage market: the twins experiment", *Journal of Political Economy* 102:113 I-1174.
- BELMAN, DALE y JOHN HEYWOOD (1991), "Sheepskin effects in the return to education", *Review of Economics and Statistics* 73: 720-724.
- BOUND, JOHN y DAVID A. JAEGER (1996), "On the validity of season of birth as an instrument in wage equations: un comentario al documento de Angrist y Krueger 'Does compulsory school attendance affect schooling and earnings?'" , Working paper no. 5835 (NBER, Cambridge, MA).
- BOUND, JOHN; DAVID A. JAEGER y REGINA M. BAKER (1995), "Problems with instrumental variables estimation when the correlation between the instruments and the endogenous explanatory variables is weak", *Journal of the American Statistical Association* 90: 443-450.
- BUTCHER, KRISTIN F. y ANNE CASE (1994), "The effect of sibling composition on women's education and earnings", *Quarterly Journal of Economics* 109:531-563.
- CAMERON, STEPHEN V. y JAMES J. HECKMAN (1993), "The non-equivalence of high school equivalents", *Journal of Labor Economics* 11: 1-47.
- CARD, DAVID (1995a), "Earnings, schooling and ability revisited", en: Solomon Polachek, editor, *Research in labor economics*, Vol. 14 (JAI Press, Greenwich, CT) pp. 23-48.
- (1995b), "Using geographic variation in college proximity to estimate the return to schooling", en: Louis N. Christofides, E. Kenneth Grant y Robert Swidinsky, editores, *Aspects of labour market behaviour: essays in honour of John Vanderkamp* (University of Toronto Press, Toronto, Canada) pp. 201-222.
- CARD, DAVID y ALAN B. KRUEGER (1992a), "Does school quality matter: returns to education and the characteristics of public schools in the United States", *Journal of Political Economy* 100: 1-40.
- (1992b), "School quality and black-white relative earnings: a direct assessment", *Quarterly Journal of Economics* 107: 151-200.
- (1996), "Labor market effects of school quality: theory and evidence", en: Gary Burtless, editores, *Does money matter? The effect of school resources on student achievement and adult success* (Brookings Institution, Washington, DC).
- CAWLEY, JOHN ; JAMES J. Heckman y EDWARD VYTLACIL (1998), "Cognitive ability and the rising return to education", Working paper no. 6388 (NBER, Cambridge, MA).
- CHAMBERLAIN, GARY (1977), "Omitted variable bias in panel data: estimating the returns to schooling", *Annals de l'Insee* 30-31: 49-82.
- (1982), "Multivariate regression models for panel data", *Journal of Econometrics* 18: 5-46.
- CHAMBERLAIN, GARY y ZVI GRILICHES (1975), "Unobservables with a variance components structure: ability, schooling and the economic success of brothers", *International Economic Review* 16: 422-449.
- (1977), "More on brothers", in: Paul Taubman, ed, *Kinometrics* (North Holland, Amsterdam).
- CHISWICK, BARRY (1974), *Income inequality: regional analysis within a human capital framework* (Columbia University Press, New York).

- COHN, ELCHANAN y JOHN T. ADDISON (1997), "The economic returns to lifelong learning", Working paper B-97- 04 (Division of Research, University of South Carolina College of Business Administration).
- CONNELLY, KAREN y ROOPE UUSITALO (1997), "Estimating heterogeneous treatment effects in the Becket schooling model", Unpublished discussion paper (Industrial Relations Section, Princeton University).
- DAHL, GORDON B. (1997), "Mobility and the returns to education: testing a Roy model with multiple markets", Unpublished discussion paper (Industrial Relations Section, Princeton University).
- DENISON, EDWARD F. (1964), "Measuring the contribution of education", in: The residual factor and economic growth (OECD, Paris).
- FREEMAN, RICHARD B. (1986), "Demand for education", en Orley Ashenfelter y Richard Layard, editores., *Handbook of labor economics* (North Holland, Amsterdam).
- GAREN, JOHN (1984), "The returns to schooling: a selectivity bias approach with a continuous choice variable", *Econometrica* 52: 1199-1218.
- GOODMAN, JERRY (1979), "The economic returns of education: an assessment of alternative models", *Social Science Quarterly* 60: 269-283.
- GRILICHES, ZVI (1970), "Notes on the role of education in production functions and growth accounting", en: W. Lee Hansen, editores, *Studies in income and wealth*, Vol. 35 (Columbia University Press, New York).
- (1977), "Estimating the returns to schooling: some econometric problems", *Econometrica* 45:1-22.
- (1979), "Sibling models and data in economics: beginnings of a survey", *Journal of Political Economy* 87: S37-S65.
- HARMON, COLIN e IAN WALKER (1995), "Estimates of the economic return to schooling for the United Kingdom", *American Economic Review* 85: 1278-1286.
- HAUSE, JOHN C. (1972), "Earnings profile: ability and schooling", *Journal of Political Economy* 80: S108-S138.
- HAUSER, ROBERT M. (1973), "Socioeconomic background and differential returns to education", en: Lewis C. Solomon y Paul J. Taubman, editores, *Does college matter? Some evidence on the impacts of higher education* (Academic Press, New York).
- HAUSMAN, JERRY A. y WILLIAM E. TAYLOR (1981), "Panel data and unobservable individual effects", *Econometrica* 49: 1377-1398.
- HECKMAN, JAMES J. y SOLOMON POLACHEK (1974), "Empirical evidence on the functional form of the earnings schooling relationship", *Journal of the American Statistical Association* 69: 350-354.
- HECKMAN, JAMES J. y EDWARD VYTLACIL (1998), "Instrumental variables methods for the correlated random coefficient model: estimating the rate of return to schooling when the return is correlated with schooling", Unpublished discussion paper (University of Chicago).
- HECKMAN, JAMES J.; ANNE LAYNE-FARRAR y PETRA TODD (1996), "Human capital pricing equations with an application to estimating the effect of school quality on earnings", *Review of Economics and Statistics* 78: 562-610.
- HUNGERFORD, THOMAS y GARY SOLON (1987), "Sheepskin effects in the return to education", *Review of Economics and Statistics* 69: 175-177.
- ICHINO, ANDREA y RUDOLF WINTER-EBMER (1998), "The long-run educational cost of World War II", Unpublished discussion paper (European University Institute).

- ISACSSON, GUNNAR (1997), "Estimates of the return to schooling in Sweden from a large sample of twins", Unpublished discussion paper (Center for Research on Transportation and Society, Borlange, Sweden).
- JAKUBSON, GEORGE (1988), "The sensitivity of labor supply parameter estimates to unobserved individual effects: fixed and random effects estimates in a nonlinear model using panel data", *Journal of Labor Economics* 6: 302-329.
- KANE, THOMAS J. y CECILIA E. ROUSE (1993), "Labor market returns to two- and four-year colleges: is a credit a credit and do degrees matter?" Working paper no. 4268 (NBER, Cambridge, MA).
- KANE, THOMAS J., CECILIA E. ROUSE y DOUGLAS STAIGER (1997), "Estimating returns to education when schooling is misreported", Unpublished discussion paper (Industrial Relations Section, Princeton University).
- KESSLER, DANIEL (1991), "Birth order, family size and achievement: family structure and wage determination", *Journal of Labor Economics* 9: 413-426.
- KOMINSKI, ROBERT y PAUL M. SIEGEL (1992), "Measuring educational attainment in the current population survey" (United States Department of Commerce Bureau of the Census, Population Division, Washington, DC).
- LANG, KEVIN (1993), "Ability bias, discount rate bias and the return to education", Unpublished discussion paper (Department of Economics, Boston University).
- LINK, CHARLES; EDWARD RATLEDGE y KENNETH LEWIS (1980), "The quality of education and cohort variation in black-white earnings differentials: reply", *American Economic Review* 70: 196-203.
- MALUCCIO, JOHN (1997), "Endogeneity of schooling in the wage function", Unpublished manuscript (Department of Economics, Yale University).
- MILLER, HERMAN P. (1955), *Income of the American people* (United States Government Printing Office, Washington, DC).
- (1966), *Income distribution in the United States* (United States Government Printing Office, Washington, DC).
- MILLER, PAUL; CHARLES MULVEY y NICK MARTIN (1995), "What do twins studies reveal about the economic returns to education? A comparison of Australian and U.S. findings", *American Economic Review* 85: 586-599.
- MINCER, JACOB (1974), *Schooling, experience and earnings* (Columbia University Press, New York).
- MURPHY, KEVIN M. y FINIS WELCH (1990), "Empirical age-earnings profiles", *Journal of Labor Economics* 8: 202-229.
- OLNECK, MICHAEL R. (1979), "The effects of education", in: Christopher Jencks *et al.*, eds., *Who gets ahead?* (Basic Books, New York).
- PARK, JIN HEUM (1994), "Returns to schooling: a peculiar deviation from linearity", Working paper no. 335 (Industrial Relations Section, Princeton University).
- (1996), "Measuring education over time", *Economics Letters* 60: 425-428.
- PSACHAROPOULOS, GEORGE (1985), "Returns to education: a further international update and implications", *Journal of Human Resources* 20: 583-604.
- (1994), "Returns to investment in education: a global update", *World Development* 22: 1325-1343.
- ROUSE, CECILIA E. (1997), "Further estimates of the economic return to schooling from a new sample of twins", Unpublished discussion paper (Industrial Relations Section, Princeton University).
- SIEBERT, W. STANLEY (1985), "Developments in the economics of human capital", en: *Labour economics* (Longman, London).

- SIEGEL, PAUL M. y ROBERT HODGE (1968), "A causal approach to the study of measurement error, en: Hubert Blalock y Ann Blalock, editores, *Methodology in social research* (McGraw Hill, New York).
- SMITH, JAMES P. y FINIS WELCH (1986), *Closing the gap: forty years of economic progress for blacks* (RAND, Santa Monica, CA).
- (1989), "Black economic progress after Myrdal", *Journal of Economic Literature* 27: 519-564.
- STAIGER, DOUGLAS y JAMES H. STOCK (1997), "Instrumental variables regression with weak instruments", *Econometrica* 65: 557-586.
- TABER, CHRISTOPHER (1998), "The college premium in the eighties: return to college or return to ability", Unpublished discussion paper (Department of Economics, Northwestern University).
- TYLER, JOHN H.; RICHARD J. MURNANE y JOHN B. WILTETT (1998), "Estimating the impact of the GED on the earnings of young dropouts using a series of natural experiments", Working paper no. 6391 (NBER, Cambridge, MA).
- WEISS, ANDREW (1995), "Human capital vs. signalling explanations of wages", *Journal of Economic Perspectives* 9: 133-154.
- WELCH, FINIS (1973), "Black-white differences in returns to schooling", *American Economic Review* 63: 893-907.
- WILLIS, ROBERT J. (1986), "Wage determinants: a survey and reinterpretation of human capital earnings functions", en: Orley Ashenfelter y Richard Layard, editores, *Handbook of labor economics* (North Holland, Amsterdam).
- WILLIS, ROBERT J. y SHERWIN ROSEN (1979), "Education and self-selection", *Journal of Political Economy* 79: S7-S36.
- WOOLDRIDGE, JEFFREY M. (1997), "On two-stage least squares estimation of the average treatment effect in a random coefficient model", *Economics Letters* 56: 129-133.
- ZHENG, JOHN (1996), "Specification testing and nonparametric estimation of the human capital model", Unpublished discussion paper (Department of Economics, University of Texas at Austin).





2

## LOS RENDIMIENTOS DE LA INVERSIÓN EN EDUCACIÓN: UNA ACTUALIZACIÓN ADICIONAL

George Psacharopoulos\*  
Harry Anthony Patrinos\*\*

---

\*G. Psacharopoulos, Universidad de Atenas, Kasou 5, Atenas 11364, Grecia. E- mail: [gpsach@erols.com](mailto:gpsach@erols.com)

\*\* H.A. Patrinos, Banco Mundial, 1818 H Street, NW, Washington DC, 20433, Estados Unidos. E-mail: [hpatri-  
nos@worldbank.org](mailto:hpatri-<br/>nos@worldbank.org)



## 2. LOS RENDIMIENTOS DE LA INVERSIÓN EN EDUCACIÓN: UNA ACTUALIZACIÓN ADICIONAL\*

### ■ RESUMEN

Los rendimientos de la inversión en educación basados en la teoría del capital humano se han estimado desde finales de los años cincuenta. En más de 40 años de historia de la estimación de los rendimientos educativos, se han realizado revisiones de resultados empíricos con el fin de establecer patrones de comportamiento. Numerosas estimaciones para una amplia variedad de países, que incluyen evidencia temporal, y basadas en nuevas técnicas econométricas, confirman la importancia de la teoría del capital humano. Este documento revisa y presenta las estimaciones más recientes, y los patrones encontrados en la literatura durante el último siglo. No obstante, debido a que la disponibilidad de estimaciones de la tasa de rendimiento ha crecido de forma exponencial, se incluye una nueva sección ante la necesidad de ser selectivos en la comparación de los rendimientos educativos y el establecimiento de patrones.

### ■ 2.1. INTRODUCCIÓN

La rentabilidad de la inversión en educación, en el moderno sentido del término “capital humano”, se ha estimado desde finales de los cincuenta. En más de 40 años de historia en las estimaciones del rendimiento de la inversión en educación, ha habido varias revisiones de resultados empíricos, con el fin de identificar patrones (véase Psacharopoulos, 1973, 1985, 1994).

El crecimiento en la desigualdad del ingreso experimentada en muchos países entre los años ochenta y noventa, renovó el interés en estimar la rentabilidad de la educación (véase, por ejemplo, Murphy y Welch, 1992). Una literatura muy extensa sugiere que cambios sistemáticos en el proceso de producción condujo a cambios en la demanda para ciertos tipos de empleo. La literatura ya había destacado que la educación debía ser más productiva cuanto más cambiante fuera el estado de la tecnología (Nelson y Phelps, 1996; Griliches, 1969; Welch, 1970; Schultz, 1975).

---

\* George Psacharopoulos y Harry Anthony Patrinos, “Returns to Investment in Education: A Further Update”, *Education Economics*, vol. 12, n.º 2, agosto 2004, páginas 111-134.

Una revisión más selectiva de las estimaciones de la tasa de rendimiento, centrándose en la discusión sobre la causalidad entre escolaridad e ingresos (Card, 2001), reafirma la conclusión de Griliches (1970) respecto a que el efecto de la habilidad y los factores relacionados no exceden el 10% del coeficiente de escolaridad estimado. Las estimaciones por variables instrumentales (VI) de los rendimientos de la educación basadas en el *background* familiar ofrecen valores más elevados que las estimaciones convencionales por Mínimos Cuadrados Ordinarios basadas en los trabajos pioneros de Mincer, Becker y Chiswick (Becker y Chiswick, 1966; Mincer, 1974). El método de estimación genera pocas diferencias en los rendimientos educativos. Las estimaciones por VI son a menudo mayores que las estimaciones por Mínimos Cuadrados Ordinarios, si bien no está claro en qué medida esto se debe a error de medida o a una instrumentación inadecuada (véase Trostel *et al.*, 2002).

En este trabajo, como es usual, comenzamos presentando las estimaciones más recientes y los patrones que de ellas se derivan. No obstante, debido a que la disponibilidad de estimaciones de la tasa de rendimiento de la educación ha crecido exponencialmente, incluimos una nueva sección atendiendo a la necesidad de ser selectivos en la comparación de los rendimientos educativos y el establecimiento de patrones.

## ■ 2.2. PATRONES RECIENTES

El patrón clásico de caída en los rendimientos educativos respecto al nivel de desarrollo económico y el nivel de escolaridad, se mantiene (véanse tablas 2.1 y 2.2 y gráficos 2.1-2.4). Por otra parte, en los datos actualizados la rentabilidad privada de la educación superior está aumentando. Estos nuevos resultados se basan en seis nuevas observaciones y la actualización de las estimaciones para 23 países desde la última revisión (Psacharopoulos,

Tabla 2.1

### RENDIMIENTOS DE LA INVERSIÓN EN EDUCACIÓN SEGÚN NIVEL, MÉTODO COMPLETO, ÚLTIMO AÑO, PROMEDIO REGIONAL (%)

REGIÓN	SOCIAL			PRIVADOS		
	PRIMARIA	SECUNDARIA	SUPERIOR	PRIMARIA	SECUNDARIA	SUPERIOR
Asia*	16,2	11,1	11,0	20,0	15,8	18,2
Europa/Oriente medio/ Norte de África*	15,6	9,7	9,9	13,8	13,6	18,8
Latinoamérica/Caribe	17,4	12,9	12,3	26,6	17,0	19,5
OECD	8,5	9,4	8,5	13,4	11,3	11,6
África Subsahariana	25,4	18,4	11,3	37,6	24,6	27,8
Mundial	18,9	13,1	10,8	26,6	17,0	19,0

Fuente: Tabla A1. \*No pertenecientes a la OECD.

Tabla 2.2

### RENDIMIENTOS DE LA INVERSIÓN EN EDUCACIÓN SEGÚN NIVEL, ÚLTIMO AÑO, PROMEDIO POR GRUPO DE INGRESO *PER CÁPITA* (Porcentaje)

GRUPO DE INGRESO PER CÁPITA	MEDIA PER CÁPITA (US\$)	SOCIAL			PRIVADOS		
		PRIMARIA	SECUNDARIA	SUPERIOR	PRIMARIA	SECUNDARIA	SUPERIOR
Ingreso alto (\$9266 o más)	22.530	13,4	10,3	9,5	25,6	12,2	12,4
Ingreso bajo (\$755 o menos)	363	21,3	15,7	11,2	25,8	19,9	26,0
Ingreso medio (\$9265)	2.996	18,8	12,9	11,3	27,4	18,0	19,3
Mundial	7.669	18,9	13,1	10,8	26,6	17,0	19,0

Fuente: Tabla A1.

1994). La información sin procesar de los rendimientos educativos para 98 países se presenta en las tablas A1-A4 del apéndice A. En estas estimaciones también se hace referencia a cuestiones de política económica. Se ha hecho un esfuerzo para seleccionar las tasas de rendimientos tan comparables como era posible (se presentan en la siguiente sección).

Los rendimientos privados son mayores que los rendimientos “sociales”, definiendo estos últimos en base a los beneficios privados y los costos totales (privados más externos) (gráfico 2.1). Esto se debe a la subvención pública de la educación, y al hecho de que la

Gráfico 2.1

### RENDIMIENTO DE LA INVERSIÓN EN EDUCACIÓN POR NIVEL, PARA EL ÚLTIMO AÑO

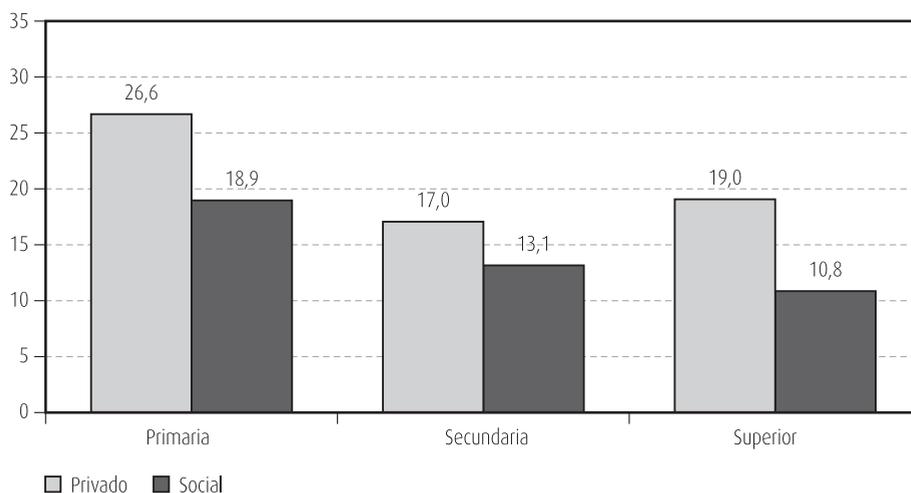


Gráfico 2.2

RENDIMIENTOS MINCERIANOS Y MEDIA DE AÑOS DE ESCOLARIDAD

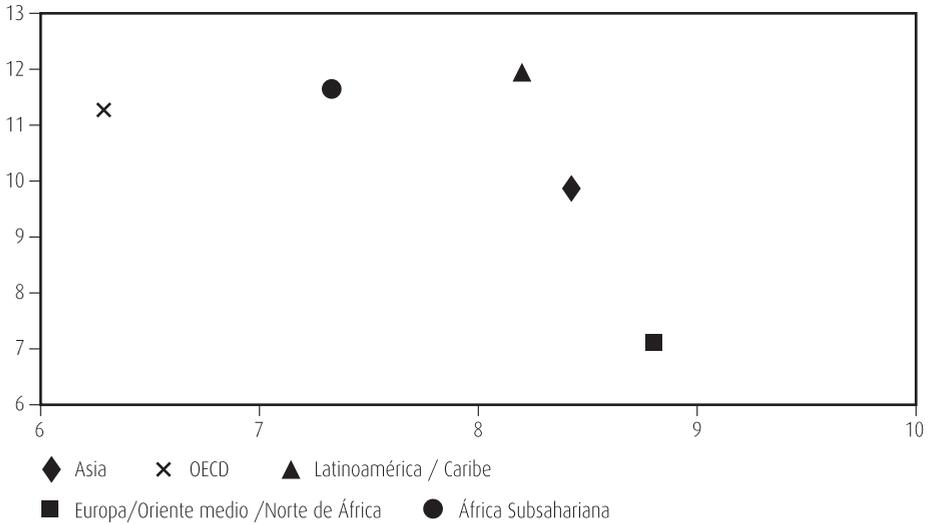


Gráfico 2.3

RENDIMIENTO SOCIAL DE LA INVERSIÓN EN EDUCACIÓN SEGÚN NIVEL DE INGRESO

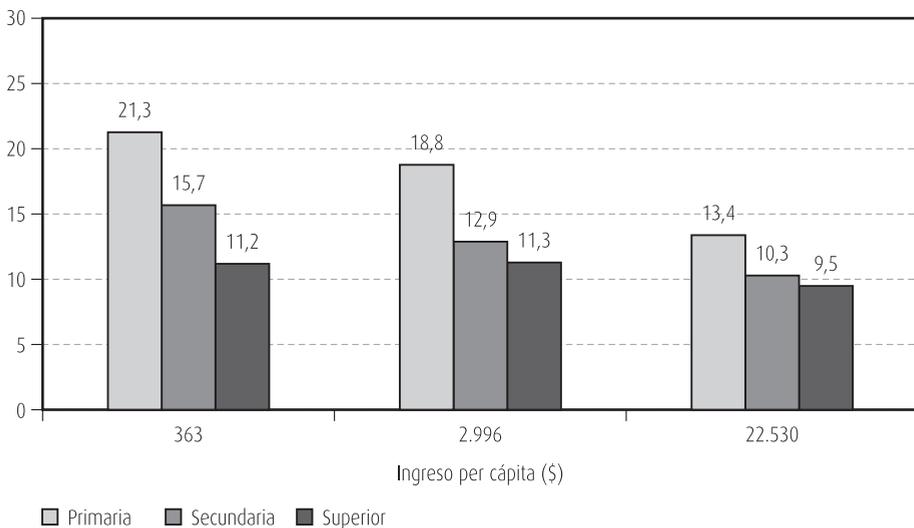
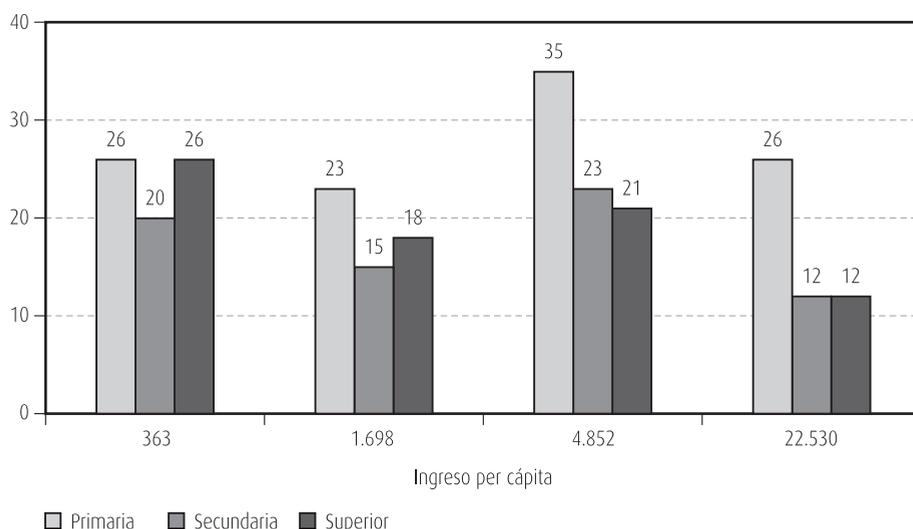


Gráfico 2.4

**RENDIMIENTO PRIVADO DE LA INVERSIÓN EN EDUCACIÓN**


estimación típica de la tasa de rendimiento social, no es capaz de incluir los beneficios sociales. No obstante, el grado de subvención pública aumenta con el nivel educativo, lo que tiene consecuencias regresivas sobre la distribución del ingreso.

En general, la tasa de rendimiento promedio de un año de educación adicional es del 10%. Los rendimientos de la educación según el nivel de ingreso del país se presentan en la tabla 2.3 (y gráfico 2.4). Los rendimientos mayores se alcanzan por los países de ingresos bajos y medios. Esta actualización incluye la estimación para nuevos países y estimaciones actualizadas para 42 países.

Tabla 2.3

**EL COEFICIENTE DE LOS AÑOS DE EDUCACIÓN: TASA DE RENDIMIENTO PROMEDIO (BASADA EN MINCER-BECKER-CHISWICK)**

GRUPO DE INGRESO PER CÁPITA	MEDIA PER CÁPITA (US\$)	AÑOS DE EDUCACIÓN	COEFICIENTE (%)
Ingreso alto (\$9266 o más)	23.463	9,4	7,4
Ingreso bajo (\$755 o menos)	375	7,6	10,9
Ingreso medio (\$9265)	3.025	8,2	10,7
Mundial	9.160	8,3	9,7

Fuente: Tabla A2.

Tabla 2.4

### EL COEFICIENTE DE LOS AÑOS DE EDUCACIÓN: TASA DE RENDIMIENTO (BASADA EN MINCER-BECKER-CHISWICK): PROMEDIO REGIONAL

REGIÓN	MEDIA PER CÁPITA (US\$)	AÑOS DE EDUCACIÓN	COEFICIENTE (%)
Asia*	5.182	8,4	9,9
Europa/Oriente medio/Norte de África*	6.299	8,8	7,1
Latinoamérica/Caribe	3.125	8,2	12,0
OECD	24.582	9,0	7,5
África Subsahariana	974	7,3	11,7
Mundial	9.160	8,3	9,7

Fuente: Tabla A2. \*No pertenecientes a la OECD.

Los mayores rendimientos educativos promedio están en la región de América Latina y el Caribe, y el África Subsahariana (tabla 2.4). Los rendimientos de la educación para Asia están alrededor de la media mundial. Los rendimientos son menores para los países de altos ingresos de la OCDE. Sorprendentemente, el menor rendimiento promedio de la educación se obtiene para los países europeos que no pertenecen a la OCDE, el Oriente medio y el grupo de países del norte de África.

Durante los últimos 12 años, el rendimiento promedio de la educación ha disminuido en 0,6 puntos porcentuales (véase la tabla A4 del Apéndice A). Al mismo tiempo, han aumentado los niveles de educación promedio. Por lo tanto, y de acuerdo con la teoría, manteniendo todo lo demás constante, un incremento en la oferta educativa ha llevado a una ligera disminución de los rendimientos educativos.

En general, las mujeres obtienen mayor rendimiento de su inversión en educación (tabla 2.5 y gráfico 2.5). Pero el rendimiento de la educación primaria es mucho mayor para

Tabla 2.5

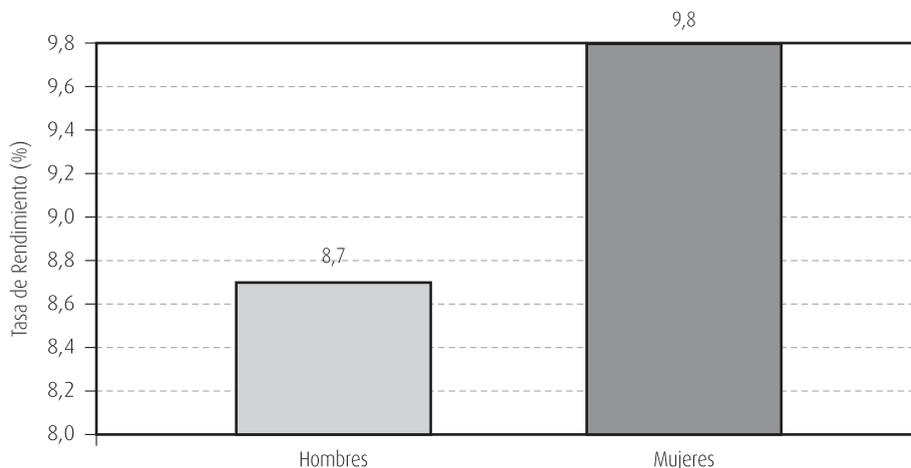
### RENDIMIENTOS DE LA EDUCACIÓN POR SEXO (Porcentaje)

NIVEL EDUCATIVO	HOMBRES	MUJERES
Primaria	20,1	12,8
Secundaria	13,9	18,4
Superior	11,0	10,8
Total	8,7	9,8

Fuente: Tabla A3.

Gráfico 2.5

## RENDIMIENTOS MINCERIANOS DE LA EDUCACIÓN POR SEXO



los hombres (20%) que para las mujeres (13%). Las mujeres, no obstante, experimentan mayores rendimientos de la educación secundaria (18% *versus* 14%).

### ■ 2.3. UN ENFOQUE MÁS SELECTIVO

Las compilaciones presentadas sobre rendimientos de la educación han sido criticadas en la literatura (véase Bennell, 1996), aunque no por razones apropiadas (véase Psacharopoulos, 1996). La razón real por la que se debería ser escéptico respecto a las compilaciones indiscriminadas sobre tasa de rendimientos, a pesar del esfuerzo de los compiladores, es que en los trabajos originales las estimaciones rara vez son plenamente comparables. Existen dos fuentes principales de incompatibilidad en la comparación: la cobertura muestral de los datos y la metodología.

Lo ideal sería, que la tasa de rendimiento de la inversión en educación estuviera basada en una muestra representativa de la población del país. Pero en la realidad, esta es más la excepción que la regla. Esto es problemático cuando las tasas de rendimiento se estiman con información de encuestas de empresas —en lugar que de hogares— ya que las muestras basadas en empresas suelen ser altamente selectivas. Con el fin de controlar el costo de estas encuestas, las muestras se centran en grandes empresas con muchos empleados. Adicionalmente, el cuestionario normalmente es rellenado por el departamento de nóminas en lugar de ser respondido por el empleado. Generalmente, este enfoque lleva al uso de muestras concentradas solo en áreas urbanas.

Otro problema surge cuando la tasa de rendimiento estimada se basa en una muestra que incluye funcionarios públicos. Esto representa un problema porque los salarios del sector público normalmente no reflejan los salarios del mercado. Por supuesto, en muchos países —aunque ahora menos que en el pasado— la mayoría de universitarios terminan en el sector público. La concentración de licenciados en el sector público se identifica como un problema en los estudios de crecimiento (véase, por ejemplo, Pissarides, 2000). Sin embargo, la tasa de rendimiento basada en base a los salarios de los funcionarios públicos, es útil en el cálculo privado respecto a los incentivos para invertir en educación —y optar por el empleo en el sector público.

Respecto a la metodología, un problema menos serio ocurre cuando el efecto salario se confunde con los rendimientos educativos. Mincer (1974) ha facilitado la estimación de los rendimientos de la educación por medio de la función semilogarítmica de ingresos (véase también Becker y Chiswick, 1966). No obstante, muchos investigadores utilizan directamente los coeficientes de la educación en la ecuación ampliada estimada por niveles educativos con variables ficticias de nivel para obtener los rendimientos de la educación cuando tales coeficientes lo que representan es el efecto salarios.

Otra limitación metodológica, a pesar de la advertencia de Becker (1964), es que los investigadores se sienten obligados a introducir cualquier variable explicativa disponible en la base de datos, entre ellas la ocupación. De hecho, este procedimiento lleva a descontar de los rendimientos de la educación la parte del efecto que proviene de la movilidad ocupacional. Por supuesto, los investigadores que incluyen ficticias de ocupación en la ecuación de ingresos lo hacen porque están interesados en modelar ingresos, no necesariamente en evaluar la tasa de rendimiento de la educación. Evidentemente, esta práctica crea un problema cuando personas diferentes de los autores de los estudios interpretan los coeficientes de escolaridad como una tasa de rendimiento Minceriana.

Quizá la estimación de los rendimientos de la educación que se detallan en el trabajo de Ashenfelter así como otros que también emplean información de gemelos u otros experimentos naturales, son los más fiables (véase Ashenfelter y Krueger, 1994; Miller *et al.*, 1995; Ashenfelter y Rouse, 1998; Rouse, 1999; Behrman y Rosenzweig, 1999). Según este trabajo, la tasa de rendimiento privada total de la inversión en educación en los Estados Unidos es aproximadamente 10%. Esta cifra establece un punto de referencia de lo que la tasa de rendimiento social debería ser (un par de puntos por debajo, si no se ajusta por externalidades), o lo que la tasa de rendimiento debería ser en un país con menor ingreso per cápita que los Estados Unidos (varios puntos porcentuales por encima, de acuerdo con la extrapolación de rendimientos educativos —no tan comparables— presentados anteriormente).

Casualmente, las estimaciones de los rendimientos educativos basadas en el análisis del ingreso de gemelos —así como las estimaciones utilizando variables instrumentales (véase, por ejemplo, Card, 2001), ascienden a una tasa promedio que es muy similar a la media mundial presentada en esta compilación: 10%.

En este estudio, sería preferible si pudiéramos incluir para cada país solo las estimaciones que incorporen la metodología más rigurosa descrita anteriormente. Esto nos ofrecería tan sólo unos pocos países. Los rendimientos presentados en las tablas A2 y A3 del Apéndice A, en su mayoría están basados en estudios sobre grandes muestras aleatorias de la población, y no se incluyen variables explicativas extrañas en la ecuación de ingresos como la ocupación, o características personales como estar casado. El sesgo de selectividad ha sido considerado para el caso de las mujeres en la mayoría de países de América Latina (basados originalmente en Psacharopoulos y Tzannatos, 1992), a pesar de que la corrección no resultó estadísticamente significativa. Pocos estudios han instrumentado los años de escolaridad.

## ■ 2.4. EXTENSIONES

Existe una discusión en la literatura respecto a lo que debería considerarse tasa de rendimiento “social”, que incluye los verdaderos beneficios sociales o externalidades. Los esfuerzos para realizar estas estimaciones son numerosos, sin embargo, los resultados varían ampliamente. Los ingresos de los individuos educados no reflejan los beneficios externos que afectan a la sociedad como un conjunto, pero que no son percibidos por el individuo. Estos beneficios son conocidos como externalidades o efectos desbordamiento, dado que se extienden sobre otros miembros de la comunidad. Son a menudo difíciles de identificar y más aún de medir. En el caso de la educación, algunos han conseguido identificar externalidades positivas pero pocos han sido capaces de cuantificarlas (véase Weisbrod, 1964; Haveman y Wolfe, 1984). Si se pudiesen incluir externalidades, entonces la tasa de rendimiento social de la educación debería ser mayor que la tasa de rendimiento privada. Una revisión reciente encontró que la evidencia empírica es escasa y poco contundente, ofreciendo algún soporte a las externalidades del capital humano, pero no muy fuertes (Venniker, 2001). Estos estudios estiman externalidades en la forma de capital humano individual que mejora la productividad de otros factores de producción a través de canales que no están interiorizados por el individuo (en el sentido de la teoría de Lucas [1988]). Como Venniker (2001) plantea, la evidencia es ambigua. De hecho, algunas estimaciones aportan valores negativos, mientras otros dan estimaciones muy altas. En países en desarrollo, nuestra revisión corrobora que la “educación para todos” impulsa, especialmente en África subsahariana. Los resultados también refuerzan el argumento de la lucha contra el trabajo infantil, no solo en términos de riesgo físico para el niño, sino en términos de pérdida de inversión.

La evidencia proviene de pocos estudios. Las regresiones entre países toman el logaritmo del producto interno bruto per capita explicado por la escolaridad media y variables de control adicionales. Los estudios micro se refieren al logaritmo del salario individual explicado por los años de escolaridad y, entre otras variables de control, los años medios de escolaridad de un área geográfica relevante. Los rendimientos sociales son iguales a la suma

de los dos coeficientes de escolaridad. Heckman y Klenow (1997) estiman la externalidad a partir de la comparación de los coeficientes de escolaridad de las regresiones entre países con los de las regresiones entre individuos. Cuando se consideran las diferencias en tecnología, los rendimientos sociales son similares a los rendimientos privados. Rauch (1993) analiza el efecto de los años medios de educación sobre el salario de los trabajadores y encuentra externalidades significativas. Sin embargo, la educación promedio y la individual pueden estar altamente correlacionadas. Acemoglu y Angrist (2000), lo corrigen utilizando variables instrumentales. Algunos estudios para África se han orientado a la estimación de los beneficios externos de la educación en la agricultura utilizando la educación de los granjeros vecinos. Un incremento de un año en el promedio de la escolaridad primaria de los granjeros vecinos, se asocia a un crecimiento del 4.3% en la producción, comparada con un efecto de 2.8% sobre la propia educación primaria de los granjeros en Uganda (Appleton y Balihuta, 1996; reportado en Appleton, 2000). Otros estudios encuentran 56% y 2% para Etiopía, pero parece realmente muy elevado (Weir, 1999; recogido en Appleton, 2000). Los resultados en general son poco concluyentes.

## ■ 2.5. ASPECTOS DE POLÍTICA

No solo la literatura académica sobre rendimientos ha aumentado, como se evidencia en este estudio, tanto en cantidad como en calidad, sino que las implicaciones de política han cambiado también. Los de la educación son interpretados como indicadores, sugiriendo áreas de investigación. Un buen ejemplo es el impacto de la tecnología sobre el diferencial de salarios, que lidera una amplia literatura sobre el cambio de la estructura salarial (véase, por ejemplo, Krueger, 1993; Patrinos, 2001).

Al mismo tiempo, la importancia de los rendimientos de la educación se hace patente en el hecho de que la OCDE (2001a) los considere un indicador clave en su serie anual de *"Panorama de la Educación"* (*Education at a Glance*) y otros documentos de política (OECD, 1997, 2001b). Cada vez más, los gobiernos y otros agentes están financiando estudios de rendimientos de la educación junto con otras investigaciones, para orientar las decisiones de política macro sobre la organización y financiación de las reformas educativas.

Los estudios sobre la tasa de rendimiento de la educación se emplean tanto para el diseño de políticas como para evaluar programas específicos. Ejemplos relevantes son el programa de construcción de escuelas en Indonesia (Duflo, 2001), el proyecto pizarra de la India (Chin, 2001) y el programa inversión en el sector principal de Etiopía (World Bank, 1998).

Sobre todo, los rendimientos educativos son útiles como indicador de productividad de la educación y como incentivo para que los individuos inviertan en su propio capital humano. Las políticas públicas necesitan considerar esta evidencia en el diseño de políticas, y generar incentivos para promover la inversión y garantizar que familias de bajos ingresos puedan beneficiarse de hacer estas inversiones.

## ■ 2.6. CONCLUSIONES

En resumen, y basados en las herramientas proporcionadas por la reciente investigación cuasi-experimental sobre economía de la educación, la inversión en educación se comporta en un sentido más o menos similar a la inversión en capital físico. En países industrialmente avanzados, los rendimientos marginales del capital humano y físico tienden a igualarse.

Al mismo tiempo, habría que destacar una importante brecha en la investigación, que es la compatibilización entre la evidencia micro y macro sobre rendimientos de la educación. Mientras que en el caso micro, como se demostró antes ampliamente, se establece más allá de cualquier duda razonable que existen rendimientos tangibles y cuantificables de la inversión en educación, dicha evidencia no es tan consistente y ni concluyente en la literatura macro (véase, por ejemplo, Pritchett, 2001; para una perspectiva diferente, véase Krueger y Lindahl, 1998; Psacharopoulos, 2000).

Se requieren más estudios sobre rendimientos sociales de la educación. Para países en desarrollo, existe una necesidad de mayor evidencia del impacto de la educación sobre los ingresos usando un diseño cuasi-experimental. Existen mayores oportunidades hoy en día para este tipo de investigación. Sin embargo, estas investigaciones deben ser utilizadas para crear programas que promuevan una mayor inversión y reformen los mecanismos de financiación.

## ■ AGRADECIMIENTOS

Las opiniones expresadas en este documento son responsabilidad de los autores y no deben ser atribuidas al Grupo del Banco Mundial. Se agradece los comentarios recibidos de Barry Chiswick, McMahon Walter, un evaluador anónimo de esta revista, y de Steve Bradley; así como las excelentes contribuciones de Georgiades Kyriakos. Los autores también agradecen la asistencia en la investigación de Leila Mamedova y Poghosyan Anahit.

## ■ BIBLIOGRAFÍA

- ACEMOGLU, D. y ANGRIST, J. (2000), *How large are human capital externalities? Evidence from compulsory schooling laws*, mimeo (Cambridge, MA, MIT).
- ALBA-RAMÍREZ, A. y SAN SEGUNDO, M.-J. (1995), "The return to education in Spain", *Economics of Education Review*, 14 (2): 155-166.
- APPLETON, S. (2000), "Education and health at the household level in sub-Saharan Africa", Working Paper No. 33 (Cambridge, MA, Harvard University Center for International Development).
- APPLETON, S. y BALIHUTA, A. (1996), "Education and agricultural productivity: evidence from Uganda", Working Paper No. WPS/96-5 (Oxford, Centre for the Study of African Economies, Oxford University).
- ARAI, M. y KJELLSTROM, C. (1999), "Returns to human capital in Sweden", en: ASPLUND, R. y PEREIRA, P. T. (eds.) *Returns to Human Capital in Europe* (Helsinki, ETLA, The Research Institute of the Finnish Economy).
- ASHENFELTER, O. y KRUEGER, A. B. (1994), "Estimates of the economic return to schooling from a new sample of twins", *American Economic Review*, 84, (5): 1157-1173.
- ASHENFELTER, O. y ROUSE, C. E. (1998), "Income, schooling, and ability: evidence from a new sample of twins", *Quarterly Journal of Economics*, 113: 253-284.
- ASPLUND, R. (1999), "Earnings and human capital: evidence for Finland", en: ASPLUND, R. y PEREIRA, P. T. (eds.) *Returns to Human Capital in Europe* (Helsinki, ETLA, The Research Institute of the Finnish Economy).
- BARTH, E. y ROED, M. (1999), "The return to human capital in Norway", en: ASPLUND, R. y PEREIRA, P. T. (eds.) *Returns to Human Capital in Europe* (Helsinki, ETLA, The Research Institute of the Finnish Economy).
- BECKER, G. S. (1964), *Human Capital: A Theoretical and Empirical Analysis* (New York, National Bureau of Economic Research).
- BECKER, G. S. y CHISWICK, B. R. (1966), "Education and the distribution of earnings", *American Economic Review*, 56: 358-369.
- BEDI, A. (1997), "The importance of school quality as determinant of earnings in a developing country: evidence from Honduras", *International Journal of Educational Development*, 17, (4): 427-437.
- BEHRMAN, J. R. y ROSENZWEIG, M. R. (1999), "'Ability' bias in schooling returns and twins: a test and new estimates", *Economics of Education Review*, 18 (2): 159-167.
- BELLI, P. y AYADI, M. A. (1998), *Returns to investment in education: the case of Nicaragua*, mimeo (Washington, DC, World Bank).
- BENNELL, P. (1996), "Rates of return to education: does the conventional pattern prevail in sub-Saharan Africa?", *World Development*, 24, (1): 183-199.
- BEVC, M. (1993), "Rates of return to investment in education in former Yugoslavia in the 1970s and 1980s by region", *Economics of Education Review*, 12, (4): 325-343.
- BRUNELLO, G.; COMI, S. y LUCIFORA, C. (1999), "Returns to education in Italy: a review of the applied literature", en: ASPLUND, R. y PEREIRA, P. T. (eds.) *Returns to Human Capital in Europe* (Helsinki, ETLA, The Research Institute of the Finnish Economy).
- CARD, D. (2001), "Estimating the return to schooling: progress on some persistent econometric problems", *Econometrica*, 69(5), pp. 1127-1160.
- CHIN, A. (2001), *The returns to school quality when school quality is very low: evidence from operation blackboard in India*, mimeo (Texas, Department of Economics, University of Houston).

- CHRISTENSEN, J. J. y WESTERGARD-NIELSEN, N. (1999), "Wages and human capital: the Danish evidence", en: ASPLUND, R. y PEREIRA, P. T. (eds.) *Returns to Human Capital in Europe* (Helsinki, ETLA, The Research Institute of the Finnish Economy).
- COHEN, B. y HOUSE, W. J. (1994), "Education, experience and earnings in the labor market of a developing economy: the case of urban Khartoum", *World Development*, 22, (10): 1549-1565.
- COHN, E. (1997), "The rate of return to schooling in Canada", *Journal of Education Finance*, 23, (2): 193-206.
- COHN, E. y ADDISON, J. T. (1998), "The economic returns to lifelong learning", *Education Economics*, 6, (3): 253-308.
- DABALEN, A. (1998), *Returns to education in Kenya and South Africa: instrumental variable estimates*, mimeo (Berkeley, CA, University of California).
- DUFOLO, E. (2001), "Schooling and labor market consequences of school construction in Indonesia: evidence from an unusual policy experiment", *American Economic Review*, 91, (4): 795-813.
- EDINVEST (1999), *The Gambia: an assessment of the private education sector*. Available online: <http://www.ifc.org/edinvest/>.
- FERSTERER, J. y WINTER-EBMER, R. (1999), "Human capital and earnings in Austria", en: ASPLUND, R. y PEREIRA, P. T. (eds.) *Returns to Human Capital in Europe* (Helsinki, ETLA, The Research Institute of the Finnish Economy).
- FILER, R., JURAJDA, S. y PLANOVSKY, J. (1999), "Education and wages in the Czech and Slovak Republics during transition", *Labour Economics*, 6: 581-593.
- FISZBEIN, A. y PSACHAROPOULOS, G. (1993), "A cost benefit analysis of educational investment in Venezuela: 1989 update", *Economics of Education Review*, 12, (4): 293-298.
- FUNKHOUSER, E. (1996), "The urban informal sector in central America", *World Development*, 24, (11): 1737-1751.
- GLEWWE, P. (1996), "The relevance of standard estimates of rates of return to schooling for education policy: a critical assessment", *Journal of Development Economics*, 51: 267-290.
- GRIFFIN, P. y COX EDWARDS, A. (1993), "Rates of return to education in Brazil: do labor market conditions matter?", *Economics of Education Review*, 12, (3): 245-256.
- GRILICHES, Z. (1969), "Capital-skill complementarity", *Review of Economics and Statistics*, 51 (4): 465-468.
- (1970), "Notes on the role of education in production functions and growth accounting", en: HANSEN, W. L. (ed.) *Education, Income and Human Capital* (New York, National Bureau of Economic Research).
- HARMON, C. y WALKER, I. (1995), "Estimates of the economic return to schooling for the United Kingdom", *American Economic Review*, 85, (5): 1278-1286.
- (1999), "The marginal and average returns to schooling in the UK", *European Economic Review*, 43: 879-887.
- HARTOG, J., ODINK, J. y SMITS, J. (1999), "Private returns to education in the Netherlands", en: ASPLUND, R. y PEREIRA, P. T. (eds.) *Returns to Human Capital in Europe* (Helsinki, ETLA, The Research Institute of the Finnish Economy).
- HAVEMAN, R. H. y WOLFE, B. (1984), "Schooling and economic well-being: the role of non-market effects", *Journal of Human Resources*, 19, (3): 128-140.
- HECKMAN, J. y KLENOW, P. (1997), *Human capital policy*, mimeo (Chicago, IL, University of Chicago).
- HOSSAIN, S. I. (1997), *Making education in China equitable and efficient*, Policy Research Working Paper No. 1814 (Washington, DC, World Bank).
- HOROWITZ, A. y SCHENZLER, C. (1999), "Returns to general, technical and vocational education in developing countries: recent evidence from Suriname", *Education Economics*, 7, (1): 5-19.

- ICHINO, A. y WINTER-EBMER, R. (1999), "Lower and upper bounds of returns to schooling: an exercise in IV estimation with different instruments", *European Economic Review*, 43: 889-901.
- ISACSSON, G. (1999), "Estimates of the return to schooling in Sweden from a large sample of twins", *Labour Economics*, 6: 471-489.
- JONES, P. (2001), "Are educated workers really more productive?", *Journal of Development Economics*, 64: 67-79.
- KATSIS, A., MATTSO, R. y PSACHAROPOULOS, G. (1998), "Explaining educational development in Pakistan: an analysis of the 1991 Household Survey", *Journal of Educational Planning and Administration*, 13(1), pp. 17-36.
- KINGDON, G. G. (1998), "Does the labour market explain lower female schooling in India?", *Journal of Development Studies*, 35, (1): 39-65.
- KLING, J. (1999), *Interpreting instrumental variable estimates of the returns of schooling*, Working Paper No. 415 (Princeton, NJ, Industrial Relations Section, Princeton University).
- KRONCKE, C. y SMITH, K. (1999), "The wage effects of ethnicity in Estonia", *Economics of Transition*, 7, (1): 179-199.
- KRUEGER, A. B. (1993), "How computers have changed the wage structure: evidence from microdata, 1984-89", *Quarterly Journal of Economics*, 108, (1): 33-60.
- KRUEGER, A. B. y LINDAHL, M. (1998), *Education for growth: why and for whom?*, mimeo (Princeton, NJ, Princeton University)
- LAMBROPOULOS, H. y KARADJIA, E. (1999), *The financing and the economics of higher education in Egypt*, mimeo (Washington, DC, World Bank).
- LASSIBILLE, G. y NAVARRO GÓMEZ, L. (1998), "The evolution of returns to education in Spain 1980-1991", *Education Economics*, 6, (1): 3-10.
- LIU, Z. (1998), "Earning, education and economic returns to urban China", *Economic Development and Cultural Change*, 46, (4): 697-725.
- LORENZ, W. y WAGNER, J. (1993), "A note on returns to human capital in the eighties: evidence from twelve countries", *Jahrbucher fur National Okonomie und Statistik*, 211, (1-2): 60-72.
- LUCAS, R. E. (1988), "On the mechanics of economic development", *Journal of Monetary Economics*, 22: 3-22.
- MAANI, S. A. (1996), "Private and social rates of return to secondary and higher education in New Zealand: evidence from the 1991 Census", *Australian Economic Review*, 1: 82-100.
- MAGOULA, T. y PSACHAROPOULOS, G. (1999), "Schooling and monetary rewards in Greece: an overeducation false alarm", *Applied Economics*, 31: 1589-1597.
- MALUCCIO, J. (1998), *Endogeneity of schooling in the wage function: evidence from the rural Philippines*, FCND Discussion Paper No. 54 (Washington, DC, International Food Policy Research Institute).
- MASON, A. y KHANDKER, S. (1997), *Household schooling decisions in Tanzania*, mimeo (Washington DC, World Bank).
- MENON, M. (1995), *Factors influencing the demand for higher education in Cyprus*, Doctoral dissertation, University of London.
- MILLER, P.; MULVEY, C. y MARTIN, N. (1995), "What do twins studies reveal about the economic returns to education? A comparison of Australian and U.S. findings", *American Economic Review*, 85, (3): 586-599.
- MINCER, J. (1974), *Schooling, Experience, and Earnings* (New York, National Bureau of Economic Research).
- MOKITIMI, N. R. y NIEUWOUDT, W. L. (1995), "Off-farm wage returns to education: evidence from Lesotho", *Development Southern Africa*, 12, (6): 839-849.

- MOOCK, P.; PATRINOS, H. y VENKATARAMAN, M. (1998), *Education and earnings in a transition economy (Vietnam)*, World Bank Policy Research Paper 1920 (Washington, DC, World Bank).
- MORA, J.-G. (1999), *Socioeconomic background, schooling, and monetary rewards in Spain*, mimeo (Valencia, University of Valencia).
- MUNICH, D.; TERRELL, K. y SVEJNAR, J. (1999), *Returns to human capital from the communist wage grid to transition: retrospective evidence from Czech micro data*, mimeo, William Davidson Institute Working Paper No. 277 (Ann Arbor, MI, University of Michigan).
- MURPHY, K. y WELCH, F. (1992), "The structure of wages", *Quarterly Journal of Economics*, 107: 285-326.
- NELSON, R. R. y PHELPS, E. S. (1966), "Investment in humans, technological diffusion, and economic growth", *American Economic Review*, 65, (2): 69-75.
- NESTEROVA, V. D. y SABIRIANOVA, K. Z. (1998), *Investing in human capital under economic transformation in Russia, Economic Education and Research Consortium*, Russia Economic Research Program, Working Paper Series No. 99/04 (Moscow).
- NOORKOIV, R.; ORAZEM, P. F.; PUUR, A. y VODOPIVEC, M. (1998), "Employment and wage dynamics in Estonia, 1989-95", *Economics of Transition*, 6, (2): 481-503.
- OECD (1997), *Human Capital Investment: An International Comparison* (Paris, OECD).
- (2001a), *Education at a Glance: OECD Indicators 2001* (Paris, OECD).
- (2001b), *Education Policy Analysis 2001* (Paris, OECD).
- PARAJULI, D. (1999), *External efficiency and equity in Nepalese education*, mimeo (Washington, DC, World Bank).
- PATRINOS, H. A. (1995), *Education and earnings differentials*, mimeo (Washington, DC, World Bank).
- (2001) *Technology and labor market outcomes: implications for education*, mimeo (Washington, DC, World Bank.)
- PATRINOS, H. A.; VELEZ, E. y PSACHAROPOULOS, G. (1994) "Language, education and earnings in Asuncion, Paraguay", *Journal of Developing Areas*, 29: 57-68.
- PISSARIDES, C. A. (2000), *Human capital and growth: a synthesis report*, OECD Development Centre, Technical Papers No. 168 (Paris, OECD).
- PRITCHETT, L. (2001), "Where has all the education gone?", *World Bank Economic Review*, 15, (3): 367-393.
- PSACHAROPOULOS, G. (1973), *Returns to Education: An International Comparison* (Amsterdam, Elsevier).
- (1985), "Returns to education: a further international update and implications", *Journal of Human Resources*, 20, (4): 583-604.
- (1994), "Returns to investment in education: a global update", *World Development*, 22 (9): 1325-1343.
- (1996), "A reply to Bennell", *World Development*, 24 (1): 201.
- (2000), *The economic costs of child labor, By the Sweat and Toil of Children* (Washington, DC, Department of Labor).
- PSACHAROPOULOS, G. y TZANNATOS, P. (1992), *Women's Employment and Pay in Latin America* (Washington, DC, World Bank).
- PSACHAROPOULOS, G. y MATTSO, R. (1998), "Estimating the returns to education: a sensitivity analysis of methods and sample size", *Journal of Educational Development and Administration*, 12, (3): 271-287.
- PSACHAROPOULOS, G.; VELEZ, E.; PANAGIDES, A. y YANG, H. (1996), "The returns to education during boom and recession: Mexico 1984, 1989 and 1992", *Education Economics*, 4, (3): 219-230.

- RAUCH, J. (1993), "Productivity gains from geographic concentration of human capital: evidence from the cities", *Journal of Urban Economics*, 34: 380-400.
- ROUSE, C. E. (1999), "Further estimates of the economic return to schooling from a new sample of twins", *Economics of Education Review*, 18, (2): 149-157.
- RUMMERY, S., VELLA, F. y VERBEEK, M. (1999), "Estimating the returns to education for Australian youth via rank-order instrumental variables", *Labour Economics*, 6: 491-507.
- RUTKOWSKI, J. (1997), "Low wage employment in transitional economies of Central and Eastern Europe", *Economic Policy and Transitional Economies*, 7, (1): 105-130.
- RYOO, J.-K.; NAM, Y.-S. y CARNOY, M. (1993), "Changing rates of return to education over time: a Korean case study", *Economics of Education Review*, 12, (1): 71-80.
- SAKELLARIOU, C. (2003), "Rates of return to investments in formal and technical/vocational education in Singapore", *Education Economics*, 11, (1): 73-87.
- SCHADY, N. R. (2000), *What education pays? Non-linear returns to schooling among Filipino men*, mimeo (Washington, DC, World Bank).
- SCHULTZ, T. W. (1975), "The value of the ability to deal with disequilibria", *Journal of Economic Literature*, 13, (3): 827-846.
- SCHULTZ, T. P. (1994), *Human capital investment in women and men*, International Center for Economic Growth Occasional Paper No. 44 (New Haven, CT, Yale University).
- TANSEL, A. (1994), "Wage employment, earning and returns to schooling for men and women in Turkey", *Economics of Education Review*, 13, (4): 305-320.
- TROSTEL, P.; WALKER, I. y WOOLLEY, P. (2002), "Estimates of the economic return to schooling for 28 countries", *Labour Economics*, 9: 1-16.
- VARGA, J. (1995), "Returns to education in Hungary", *Acta Oeconomica*, 47, (1-2): 201-216.
- VENNIKER, R. (2001), Social returns to education: a survey of recent literature on human capital externalities, CPB (Netherlands Bureau for Economic Policy Analysis) Report 00/1. Available online: ([http://www.cpb.nl/eng/cpbreport/2000\\_1/s3\\_4.pdf](http://www.cpb.nl/eng/cpbreport/2000_1/s3_4.pdf)).
- WEBER, B. y WOLTER, S. (1999), "Wages and human capital: evidence from Switzerland", en: ASPLUND, R. y PEREIRA, P. T. (eds.) *Returns to Human Capital in Europe* (Helsinki, ETLA, The Research Institute of the Finnish Economy).
- WEI, X., TSANG, M. C., WU, W. y CHEN, L.-K. (1999), "Education and earning in rural China", *Education Economics*, 7, (2): 167-188.
- WEIR, S. (1999), The effects of education on farmer productivity in rural Ethiopia, Centre for the Study of African Economies Working Paper No. WPS/99-7 (Oxford, Oxford University).
- WEISBROD, B. A. (1964), *External Benefits of Education* (Princeton, NJ, Princeton University, Industrial Relations Section).
- WELCH, F. (1970), "Education in production", *Journal of Political Economy*, 78, (1): 35-59.
- WORLD BANK (1996a), *Republic of Ghana: basic education sector improvement program*, Report No. 15570-GH (Washington, DC, World Bank).
- (1996b), *Bolivia: Poverty, Equity, and Income*, Report No. 15272-BO (Washington, DC, World Bank).
- (1998), *Ethiopia: Education Sector Development Program*, Report No. 17739-ET. (Washington, DC, World Bank).

## ■ APÉNDICE A

Tabla A.1

**RENDIMIENTOS DE LA INVERSIÓN EN EDUCACIÓN POR NIVEL  
(PORCENTAJE), MÉTODO COMPLETO, ÚLTIMO AÑO**

PAÍS	AÑO	SOCIAL			PRIVADO			FUENTE
		PRIMARIA	SECUNDARIA	SUPERIOR	PRIMARIA	SECUNDARIA	SUPERIOR	
Argentina	1989	8,4	7,1	7,6	10,1	14,2	14,9	Psacharopoulos (1994)
Australia	1976			16,3		8,1	21,1	Psacharopoulos (1994)
Austria	1981					11,3	4,2	Psacharopoulos (1994)
Bahamas	1970			20,6		26,1		Psacharopoulos (1994)
Bélgica	1960			17,1	6,7	21,2	8,7	Psacharopoulos (1994)
Bolivia	1990	13,0	6,0	13,0	20,0	6,0	19,0	Psacharopoulos (1994)
Botswana	1983	42,0	41,0	15,0	99,0	76,0	38,0	Psacharopoulos (1994)
Brasil	1989	35,6	5,1	21,4	36,6	5,1	28,2	Psacharopoulos (1994)
Burkina Faso	1982	20,1	14,9	21,3				Psacharopoulos (1994)
Canadá	1994					7,8	13,0	Cohn (1997)
Chile	1989	8,1	11,1	14,0	9,7	12,9	20,7	Psacharopoulos (1994)
China	1993	14,4	12,9	11,3	18,0	13,4	15,1	Hossain (1997)
Colombia	1989	20,0	11,4	14,0	27,7	14,7	21,7	Psacharopoulos (1994)
Costa Rica	1989	11,2	14,4	9,0	12,2	17,6	12,9	Psacharopoulos (1994)
Chipre	1979	7,7	6,8	7,6	15,4	7,0	5,6	Psacharopoulos (1994)
Dinamarca	1964			7,8			10,0	Psacharopoulos (1994)
Dominicana República	1989				85,1	15,1	19 A	Psacharopoulos (1994)
Ecuador	1987	14,7	12,7	9,9	17,1	17,2	12,7	Psacharopoulos (1994)
El Salvador	1990	16,4	13,3	8,0	18,9	14,5	9,5	Psacharopoulos (1994)
Estonia	1995	14,0	2,2	10,3				Noorkoiv <i>et al.</i> , (1998)
Etiopía	1996	14,9	14,4	11,9	24,7	24,2	26,6	World Bank (1998)
Francia	1976					14,8	20,0	Psacharopoulos (1994)
Alemania (Occidental)	1978					6,5	10,5	Psacharopoulos (1994)
Ghana	1967	18,0	13,0	16,5	24,5	17,0	37,0	Psacharopoulos (1994)
Grecia	1993		6,5	5,7		8,3	8,1	Magoula y Psacharopoulos (1999)
Guatemala	1989				33,8	17,9	22,2	Psacharopoulos (1994)
Honduras	1989	18,2	19,7	18,9	20,8	23,3	25,9	Psacharopoulos (1994)
Hong Kong	1976		15,0	12,4		18,5	25,2	Psacharopoulos (1994)
Hungría	1993		6,0	2,0		8,2	13,4	Varga (1985)

Tabla A.1 (continuación)

## RENDIMIENTOS DE LA INVERSIÓN EN EDUCACIÓN POR NIVEL (PORCENTAJE), MÉTODO COMPLETO, ÚLTIMO AÑO

PAÍS	AÑO	SOCIAL			PRIVADO			FUENTE
		PRIMARIA	SECUNDARIA	SUPERIOR	PRIMARIA	SECUNDARIA	SUPERIOR	
India	1995				2,6	17,6	18,2	Kingdon (1998)
Indonesia	1989		11,0	5,0		21,2	18,5	Psacharopoulos (1994)
Irán	1976	15,2	17,6	13,6				Psacharopoulos (1994)
Israel	1958	16,5	6,9	6,6	27,0	6,9	8,0	Psacharopoulos (1994)
Italia	1969					17,3	18,3	Psacharopoulos (1994)
Costa de Marfil	1984				25,7	30,7	25,1	Psacharopoulos (1994)
Jamaica	1989	17,7	7,9		20,4	15,7		Psacharopoulos (1994)
Japón	1976	9,6	8,6	6,9	13,4	10,4	8,8	Psacharopoulos (1994)
Kenia	1980		10,0			16,0		Psacharopoulos (1994)
Corea	1986		8,8	15,5		10,1	17,9	Psacharopoulos (1994)
Lesotho	1980	10,7	18,6	10,2	15,5	26,7	36,5	Psacharopoulos (1994)
Liberia	1983	41,0	17,0	8,0	99,0	30,5	17,0	Psacharopoulos (1994)
Malawi	1982	14,7	15,2	11,5	15,7	16,8	46,6	Psacharopoulos (1994)
Malasia	1978					32,6	34,5	Psacharopoulos (1994)
México	1992	11,8	14,6	11,1	18,9	20,1	15,7	Cohn y Addison (1998)
Marruecos	1970	50,5	10,0	13,0				Psacharopoulos (1994)
Nepal	1999	15,7	8,1	9,1	16,6	8,5	12,0	Parajuli (1999)
Holanda	1965		5,2	5,5		8,5	10,4	Psacharopoulos (1994)
Nueva Zelanda	1991		12,4	9,5		13,8	11,9	Maani (1996)
Nicaragua	1996	13,6	10,4	14,7				Belli y Ayadi (1998)
Nigeria	1966	23,0	12,8	17,0	30,0	14,0	34,0	Psacharopoulos (1994)
Noruega	1966		7,2	7,5		7,4	7,7	Psacharopoulos (1994)
Pakistán	1991				8,4	13,7	31,2	Katsis <i>et al.</i> , (1998)
Panamá	1989				5,7	21,0	21,0	Psacharopoulos (1994)
Papua Nueva Guinea	1986	12,8	19,4	8,4	37,2	41,6	23,0	Psacharopoulos (1994)
Paraguay	1990	20,3	12,7	10,8	23,7	14,6	13,7	Psacharopoulos (1994)
Perú	1990				13,2	6,6	40,0	Psacharopoulos (1994)
Filipinas	1988	13,3	8,9	10,5	18,3	10,5	11,6	Psacharopoulos (1994)
Puerto Rico	1959	24,0	34,1	15,5	68,2	52,1	29,0	Psacharopoulos (1994)
Senegal	1985	23,0	8,9		33,7	21,3		Psacharopoulos (1994)
Sierra Leona	1971	20,0	22,0	9,5				Psacharopoulos (1994)
Singapur	1998	16,7	10,1	13,9	22,2	12,9	18,7	Sakellariou (2003)
Somalia	1983	20,6	10,4	19,9	59,9	13,0	33,2	Psacharopoulos (1994)

Tabla A.1 (continuación)

### RENDIMIENTOS DE LA INVERSIÓN EN EDUCACIÓN POR NIVEL (PORCENTAJE), MÉTODO COMPLETO, ÚLTIMO AÑO

PAÍS	AÑO	SOCIAL			PRIVADO			FUENTE
		PRIMARIA	SECUNDARIA	SUPERIOR	PRIMARIA	SECUNDARIA	SUPERIOR	
Sudáfrica	1980	22,1	17,7	11,8				Psacharopoulos (1994)
España	1991	7,4	8,5	13,5				Lassibille y Navarro (1998)
Sri Lanka	1981					12,6	16,1	Psacharopoulos (1994)
Sudan	1974		8,0	4,0		13,0	15,0	Psacharopoulos (1994)
Suecia	1967		10,5	9,2			10,3	Psacharopoulos (1994)
Taiwan	1972	27,0	12,3	17,7	50,0	12,7	15,8	Psacharopoulos (1994)
Tanzania	1991					7,9	8,5	Masón y Khandker (1997)
Tailandia	1989				16,0	12,9	11,8	Schultz (1994)
Gambia	1997	33,5	12,1		37,1	12,7		EdInvest (1999)
Túnez	1980					13,0	27,0	Psacharopoulos (1994)
Turquía	1987			8,5	1,0	8,6	16,2	Tansel (1994)
Uganda	1965	66,0	28,6	12,0				Psacharopoulos (1994)
Reino Unido	1986	8,6	7,5	6,5				Cohn y Addison (1998)
Estados Unidos	1987		10,0	12,0				Psacharopoulos (1994)
Uruguay	1989	21,6	8,1	10,3	27,8	10,3	12,8	Psacharopoulos (1994)
Venezuela	1989	23,4	10,2	6,2	36,3	14,6	11,0	Psacharopoulos (1994)
Vietnam	1992	13,5	4,5	6,2	10,8	3,8	3,0	Mooock <i>et al.</i> (1998)
Yemen	1985	2,0	26,0	24,0	10,0	41,0	56,0	Psacharopoulos (1994)
Yugoslavia	1986	3,3	2,3	3,1	14,6	3,1	5,3	Psacharopoulos (1994)
Zambia	1983			5,7			19,2	Psacharopoulos (1994)
Zimbabwe	1987	11,2	47,6	-4,3	16,6	48,5	5,1	Psacharopoulos (1994)

Tabla A.2

## EL COEFICIENTE DE LOS AÑOS DE EDUCACIÓN, ÚLTIMO AÑO

PAÍS	AÑO	AÑOS DE ESCOLARIDAD PROMEDIO	COEFICIENTE (%)	FUENTE
Argentina	1989	9,1	10,3	Psacharopoulos (1994)
Australia	1989		8,0	Cohn y Addison (1998)
Austria	1993		7,2	Fersterer y Winter-Ebmer (1999)
Bolivia	1993		10,7	Patrinos (1995)
Botswana	1979	3,3	19,1	Psacharopoulos (1994)
Brasil	1989	5,3	14,7	Psacharopoulos (1994)
Burkina Faso	1980		9,6	Psacharopoulos (1994)
Canadá	1989		8,9	Cohn (1997)
Chile	1989	8,5	12,0	Psacharopoulos (1994)
China	1993		12,2	Hossain(1997)
Colombia	1989	8,2	14,0	Psacharopoulos (1994)
Costa Rica	1991		8,5	Funkhouser (1996)
Costa de Marfil	1986	6,9	20,1	Psacharopoulos (1994)
Chipre	1994		5,2	Menon (1995)
Dinamarca	1990		4,5	Christensen y Westergard-Nielsen (1999)
Dominicana República	1989	8,8	9,4	Psacharopoulos (1994)
Ecuador	1987	9,6	11,8	Psacharopoulos (1994)
Egipto	1997		5,2	Lambropoulos y Karadjia (1999)
El Salvador	1992		7,6	Funkhouser (1996)
Estonia	1994	10,9	5,4	Kroncke y Smith (1999)
Etiopía	1972	6,0	8,0	Psacharopoulos (1994)
Finlandia	1993		8,2	Asplund (1999)
Francia	1977	6,2	10,0	Psacharopoulos (1994)
Alemania	1988		7,7	Cohn y Addison (1998)
Ghana	1995	9,7	7,1	Jones (2001)
Grecia	1993		7,6	Magoula y Psacharopoulos (1999)
Guatemala	1989	4,3	14,9	Psacharopoulos (1994)
Honduras	1991		9,3	Funkhouser (1996)
Hong Kong	1981	9,1	6,1	Psacharopoulos (1994)
Hungría	1987	11,3	4,3	Psacharopoulos (1994)
India	1995		10,6	Kingdon (1998)
Indonesia	1995	8,0	7,0	Duflo (2001)
Irán	1975		11,6	Psacharopoulos (1994)
Israel	1979	11,2	6,4	Psacharopoulos (1994)
Italia	1987		2,7	Brunello <i>et al.</i> (1999)

Tabla A.2 (continuación)

## EL COEFICIENTE DE LOS AÑOS DE EDUCACIÓN, ÚLTIMO AÑO

PAÍS	AÑO	AÑOS DE ESCOLARIDAD PROMEDIO	COEFICIENTE (%)	FUENTE
Jamaica	1989	7,2	28,8	Psacharopoulos (1994)
Japón	1988		13,2	Cohn y Addison (1998)
Kenia	1986	8,0	16,0	Dabalen (1998)
Corea	1986	8,0	13,5	Ryoo <i>et al.</i> (1993)
Kuwait	1983	8,9	4,5	Psacharopoulos (1994)
Malasia	1979	15,8	9,4	Psacharopoulos (1994)
México	1992		7,6	Psacharopoulos <i>et al.</i> (1996)
Marruecos	1970	2,9	15,8	Psacharopoulos (1994)
Nepal	1999	3,9	9,7	Parajuli (1999)
Holanda	1994		6,4	Hartog <i>et al.</i> (1999)
Nicaragua	1996		12,1	Belli y Ayadi (1998)
Noruega	1995		5,5	Barth y Roed (1999)
Pakistán	1991		15,4	Katsis <i>et al.</i> (1998)
Panamá	1990	9,2	13,7	Psacharopoulos (1994)
Paraguay	1990	9,1	11,5	Psacharopoulos (1994)
Perú	1990	10,1	8,1	Psacharopoulos (1994)
Filipinas	1998	8,8	12,6	Schady (2000)
Polonia	1995-1996		7,0	Nesterova y Sabirianova (1998)
Portugal	1991		8,6	Cohn y Addison (1998)
Puerto Rico	1989		15,1	Griffin y Cox Edwards (1993)
Federación de Rusia	1996	11,7	7,2	Nesterova y Sabirianova (1998)
Singapur	1998	9,5	13,1	Sakellariou (2003)
Sudáfrica	1994	7,1	4,1	Dabalen (1998)
España	1991		7,2	Mora (1999)
Sri Lanka	1981	4,5	7,0	Psacharopoulos (1994)
Sudan	1989	10,2	9,3	Cohen y House (1994)
Suecia	1991		5,0	Cohn y Addison (1998)
Suiza	1991		7-5	Weber y Wolter (1999)
Taiwan	1972	9,0	6,0	Psacharopoulos (1994)
Tanzania	1980		11,9	Psacharopoulos (1994)
Tailandia	1989		11,5	Patrinos (1995)
Túnez	1980	4,8	8,0	Psacharopoulos (1994)
Reino Unido	1987	11,8	6,8	Psacharopoulos (1994)
Estados Unidos	1991-1995		10,0	Rouse (1999)
Uruguay	1989	9,0	9,7	Psacharopoulos (1994)
Venezuela	1992		9,4	Psacharopoulos y Mattson (1998)
Vietnam	1992	7,9	4,5	Mooch <i>et al.</i> (1998)
Yugoslava	1986		4,8	Bevc (1993)

Tabla A.3

## RENDIMIENTOS DE LA EDUCACIÓN POR NIVEL EDUCATIVO Y GÉNERO

PAÍS	EDUCACIÓN			FUENTE
	AÑO	NIVEL	HOMBRES MUJERES	
Argentina	1985	Total	9,1 10,3	Psacharopoulos (1994)
Argentina	1989	Total	10,7 11,2	Psacharopoulos (1994)
Austria	1981	Total	10,3 13,5	Psacharopoulos (1994)
Bolivia	1989	Total	7,3 7,7	Psacharopoulos (1994)
Botswana	1975	Total	16,4 18,2	Psacharopoulos (1994)
Brasil	1980	Total	14,7 15,6	Psacharopoulos (1994)
Brasil	1989	Total	15,4 14,2	Psacharopoulos (1994)
Canadá	1989	Total	8,9	Patrinos (1995)
República Checa	1984	Total	2,4 4,2	Munich <i>et al.</i> (1999)
República Checa	1988	Total	4,0 5,7	Nesterova y Sabirianova (1998)
República Checa	1989	Total	2,7 3,8	Munich <i>et al.</i> (1999)
República Checa	1992	Total	5,3 6,7	Nesterova y Sabirianova (1998)
República Checa	1993	Total	5,2 5,8	Munich <i>et al.</i> (1999)
República Checa	1996	Total	5,8 7,0	Munich <i>et al.</i> (1999)
Chile	1987	Total	13,7 12,6	Psacharopoulos (1994)
Chile	1989	Total	12,1 13,2	Psacharopoulos (1994)
China	1985	Total	4,5 5,6	Psacharopoulos (1994)
Colombia	1973	Total	18,1 20,8	Psacharopoulos (1994)
Colombia	1973	Total	10,3 20,1	Psacharopoulos (1994)
Colombia	1988	Total	11,1 9,7	Psacharopoulos (1994)
Colombia	1989	Total	14,5 12,9	Psacharopoulos (1994)
Costa Rica	1974	Total	14,7 14,7	Psacharopoulos (1994)
Costa Rica	1989	Total	10,1 13,1	Psacharopoulos (1994)
Costa Rica	1989	Total	10,5 13,5	Psacharopoulos (1994)
Costa de Marfil	1987	Total	13,6 12,1	Schultz (1994)
Chipre	1984	Total	8,9 12,7	Psacharopoulos (1994)
Dinamarca	1990	Total	5,1 3,4	Christensen y Westergard-Nielsen(1999)
República Dominicana	1989	Total	7,8 12,0	Psacharopoulos (1994)
Ecuador	1987	Total	11,4 10,7	Psacharopoulos (1994)
Ecuador	1987	Total	9,8 11,5	Psacharopoulos (1994)
El Salvador	1990	Total	9,6 9,8	Psacharopoulos (1994)
Finlandia	1980	Total	9,3 -	Asplund (1999)
Finlandia	1987	Total	7,4 6,4	Asplund(1999)
Finlandia	1989	Total	8,4 7,8	Asplund (1999)

Tabla A.3 (continuación)

## RENDIMIENTOS DE LA EDUCACIÓN POR NIVEL EDUCATIVO Y GÉNERO

PAÍS	EDUCACIÓN				FUENTE
	AÑO	NIVEL	HOMBRES	MUJERES	
Finlandia	1991	Total	8,8	8,7	Asplund (1999)
Finlandia	1993	Total	7,8	8,3	Asplund (1999)
Alemania	1974	Total	13,1	11,2	Psacharopoulos (1994)
Alemania	1977	Total	13,6	11,7	Psacharopoulos (1994)
Ghana	1989	Total	4,4	4,2	Schultz (1994)
Ghana	1992	Total	9,3	10,6	World Bank (1996a)
Grecia	1977	Total	4,7	4,5	Psacharopoulos (1994)
Grecia	1993	Total	6,9	9,0	Magoula y Psacharopoulos (1999)
Guatemala	1989	Total	14,2	16,3	Psacharopoulos (1994)
Honduras	1989	Total	17,2	19,8	Psacharopoulos (1994)
India	1978	Total	5,3	3,6	Psacharopoulos (1994)
Italia	1978	Total	2,9	—	Brunello <i>et al.</i> (1999)
Italia	1983	Total	6,0	3,5	Brunello <i>et al.</i> (1999)
Italia	1985	Total	3,5	3,9	Brunello <i>et al.</i> (1999)
Italia	1987	Total	0,0	3,0	Brunello <i>et al.</i> (1999)
Costa de Marfil	1984	Total	11,1	22,6	Psacharopoulos (1994)
Jamaica	1989	Total	12,3	21,5	Psacharopoulos (1994)
Jamaica	1989	Total	28,0	31,7	Psacharopoulos (1994)
Lesotho	1993	Total	10,6	16,5	Mokitimi y Nieuwoudt (1995)
Malasia	1979	Total	5,3	8,2	Psacharopoulos (1994)
México	1984	Total	13,2	14,7	Psacharopoulos (1994)
México	1984	Total	14,1	15,0	Psacharopoulos (1994)
Nicaragua	1978	Total	8,5	11,5	Psacharopoulos (1994)
Noruega	1980	Total	3,2	4,9	Barth y Roed (1999)
Noruega	1983	Total	4,7	5,5	Banh y Roed (1999)
Noruega	1987	Total	4,3	6,2	Barth y Roed (1999)
Noruega	1989	Total	4,1	3,9	Barth y Roed (1999)
Noruega	1991	Total	4,2	5,3	Barth y Roed (1999)
Panamá	1989	Total	9,7	11,9	Psacharopoulos (1994)
Panamá	1989	Total	12,6	17,1	Psacharopoulos (1994)
Paraguay	1990	Total	10,3	12,1	Psacharopoulos (1994)
Perú	1985	Total	11,5	12,4	Psacharopoulos (1994)
Perú	1990	Total	8,5	6,5	Psacharopoulos (1994)
Filipinas	1988	Total	12,4	12,4	Psacharopoulos (1994)
Portugal	1977	Total	7,5	8,4	Psacharopoulos (1994)

Tabla A.3 (continuación)

## RENDIMIENTOS DE LA EDUCACIÓN POR NIVEL EDUCATIVO Y GÉNERO

PAÍS	EDUCACIÓN				FUENTE
	AÑO	NIVEL	HOMBRES	MUJERES	
Portugal	1985	Total	9,4	10,4	Psacharopoulos (1994)
Rusia	1991	Total	3,1	5,4	Munich et al, (1999)
Rusia	1994	Total	6,7	9,6	Munich et al, (1999)
Singapur	1998	Total	11,1	15,2	Sakellariou (2003)
Eslovaquia	1984	Total	2,8	4,4	Munich et al, (1999)
Eslovaquia	1993	Total	4,9	5,4	Munich et al, (1999)
Sudáfrica	1990	Total	2,3	0,8	Mokitimi y Nieuwoudt (1995)
Corea del Sur	1976	Total	10,3	1,7	Psacharopoulos (1994)
Corea del Sur	1980	Total	17,2	5,0	Psacharopoulos (1994)
Sri Lanka	1981	Total	6,9	7,9	Psacharopoulos (1994)
Suecia	1968	Total	8,9	8,7	Arai y Kjellstrom (1999)
Suecia	1974	Total	5,3	5,5	Arai y Kjellstrom (1999)
Suecia	1980	Total	4,3	4,3	Arai y Kjellstrom (1999)
Suecia	1981	Total	5,1	4,1	Arai y Kjellstrom (1999)
Suecia	1991	Total	5,0	4,0	Arai y Kjellstrom (1999)
Suiza	1982	Total	5,5	9,1	Weber y Wolter (1999)
Suiza	1991	Total	8,3	7,5	Weber y Wolter (1999)
Suiza	1992	Total	8,2	7,8	Weber y Wolter (1999)
Suiza	1993	Total	7,8	7,9	Weber y Wolter (1999)
Suiza	1995	Total	9,1	9,0	Weber y Wolter (1999)
Suiza	1997	Total	—	6,1	Weber y Wolter (1999)
Tailandia	1972	Total	9,1	13,0	Psacharopoulos (1994)
Uruguay	1989	Total	9,0	10,6	Psacharopoulos (1994)
Venezuela	1976	Total	9,9	13,5	Psacharopoulos (1994)
Venezuela	1987	Total	10,0	13,1	Psacharopoulos (1994)
Venezuela	1989	Total	9,1	11,1	Psacharopoulos (1994)
Venezuela	1989	Total	8,4	8,0	Psacharopoulos (1994)
Vietnam	1992	Total	3,4	6,8	Mooock et al, (1998)
Yugoslavia	1976	Total	5,8	6,6	Psacharopoulos (1994)
Yugoslavia	1986	Total	4,9	4,8	Psacharopoulos (1994)
<b>Media</b>			<b>8,7</b>	<b>9,8</b>	
Puerto Rico	1959	Primaria	29,5	18,4	Psacharopoulos (1994)
Taiwan	1982	Primaria	8,4	16,1	Psacharopoulos (1994)
Indonesia	1982	Primaria	19,0	17,0	Psacharopoulos (1994)
Gran Bretaña	1841	Alfabetismo	24,5	3,5	Psacharopoulos (1994)

Tabla A.3 (continuación)

## RENDIMIENTOS DE LA EDUCACIÓN POR NIVEL EDUCATIVO Y GÉNERO

PAÍS	EDUCACIÓN				FUENTE
	AÑO	NIVEL	HOMBRES	MUJERES	
Gran Bretaña	1871	Alfabetismo	19,0	9,0	Psacharopoulos (1994)
<b>Media</b>			<b>20,1</b>	<b>12,8</b>	
Canadá	1980	Secundaria	2,0	6,0	Psacharopoulos (1994)
Canadá	1985	Secundaria	10,6	18,6	Psacharopoulos (1994)
Francia	1969	Secundaria	13,9	15,4	Psacharopoulos (1994)
Francia	1976	Secundaria	14,8	16,2	Psacharopoulos (1994)
Gran Bretaña	1971	Secundaria	10,0	8,0	Psacharopoulos (1994)
Indonesia	1982	Secundaria	23,0	11,0	Psacharopoulos (1994)
Indonesia	1986	Secundaria	11,0	16,0	Psacharopoulos (1994)
Puerto Rico	1959	Secundaria	27,3	40,8	Psacharopoulos (1994)
Corea del Sur	1971	Secundaria	13,7	16,9	Psacharopoulos (1994)
Sri Lanka	1981	Secundaria	12,6	35,5	Psacharopoulos (1994)
Suriname	1993	Secundaria	10,7	-0,8	Horowitz y Schenzler (1999)
<b>Media</b>			<b>13,9</b>	<b>18,4</b>	
Australia	1976	Universidad	21,1	21,2	Psacharopoulos (1994)
Canadá	1980	Universidad	5,5	10,5	Psacharopoulos (1994)
Canadá	1985	Universidad	8,3	18,8	Psacharopoulos (1994)
Chipre	1994	Superior	5,2	7,2	Menon (1995)
Dinamarca	1990	Universidad	3,5	5,2	Cohn y Addison (1998)
Finlandia	1987	Universidad	6,6	7,7	Cohn y Addison (1998)
Francia	1969	Universidad	22,5	13,8	Psacharopoulos (1994)
Francia	1976	Universidad	20,0	12,7	Psacharopoulos (1994)
Francia	1976	Universidad	20,0	12,7	Psacharopoulos (1994)
Gran Bretaña	1971	Universidad	8,0	12,0	Psacharopoulos (1994)
Indonesia	1982	Universidad	10,0	9,0	Psacharopoulos (1994)
Indonesia	1986	Universidad	9,0	10,0	Psacharopoulos (1994)
Japón	1976	Universidad	6,9	6,9	Psacharopoulos (1994)
Japón	1980	Universidad	5,7	5,8	Psacharopoulos (1994)
Noruega	1991	Universidad	4,0	4,2	Cohn y Addison (1998)
Puerto Rico	1959	Universidad	21,9	9,0	Psacharopoulos (1994)
Corea del Sur	1971	Universidad	15,7	22,9	Psacharopoulos (1994)
Suecia	1991	Universidad	4,4	5,0	Cohn y Addison (1998)
<b>Media</b>			<b>11,0</b>	<b>10,8</b>	

Tabla A.4

### RENDIMIENTOS DE LA INVERSIÓN EN EDUCACIÓN COMPARABLES EN EL TIEMPO EN PAÍSES SELECCIONADOS

PAÍS	AÑO	TASA DE RENDIMIENTO (%)	FUENTE
Australia	1980	7,9	Millar <i>et al.</i> (1995)
Australia	1981	8,4	Patrinos (1995)
Australia	1985	10,9	Rummery <i>et al.</i> (1999)
Australia	1987	5,4	Patrinos (1995)
Australia	1989	8,0	Cohn y Addison (1998)
Austria	1981	11,6	Psacharopoulos (1994)
Austria	1983	7,9*	Fersterer y Winter-Ebmer (1999)
Austria	1985	7,6*	Fersterer y Winter-Ebmer (1999)
Austria	1987	7,4*	Fersterer y Winter-Ebmer (1999)
Austria	1989	7,6*	Fersterer y Winter-Ebmer (1999)
Austria	1991	7,4*	Fersterer y Winter-Ebmer (1999)
Austria	1993	7,2*	Fersterer y Winter-Ebmer (1999)
Bolivia	1981	12,2	World Bank (1996b)
Bolivia	1988	9,5	World Bank (1996b)
Bolivia	1989	7,1	World Bank (1996b)
Bolivia	1990	10,i	Psacharopoulos y Mattson (1998)
Bolivia	1991	8,9	World Bank (1996b)
Bolivia	1992	10,0	World Bank (1996b)
Bolivia	1993	10,7	Patrinos (1995)
Brasil	1970	16,5	Patrinos (1995)
Brasil	1980	14,5	Patrinos (1995)
Brasil	1989	14,7	Patrinos (1995)
Canadá	1981	8,5	Patrinos (1995)
Canadá	1986	8,8	Patrinos (1995)
Canadá	1989	8,9	Patrinos (1995)
Chile	1974	8,3	Patrinos (1995)
Chile	1980	9,6	Patrinos (1995)
Chile	1989	8,2	Patrinos (1995)
China	1988	3,6	Liu (1998)
China	1991	4,8	Wei <i>et al.</i> (1999)
China	1993	12,2	Hossain (1997)
Colombia	1965	17,3	Patrinos (1995)
Colombia	1974	12,8	Patrinos (1995)
Colombia	1989	8,2	Patrinos (1995)
Costa Rica	1980	10,5	Funkhouser (1996)

Tabla A.4 (continuación)

### RENDIMIENTOS DE LA INVERSIÓN EN EDUCACIÓN COMPARABLES EN EL TIEMPO EN PAÍSES SELECCIONADOS

PAÍS	AÑO	TASA DE RENDIMIENTO (%)	FUENTE
Costa Rica	1983	8,1	Funkhouser (1996)
Costa Rica	1985	8,1	Funkhouser (1996)
Costa Rica	1988	9,1	Funkhouser (1996)
Costa Rica	1991	8,5	Funkhouser (1996)
Chipre	1984	11,0	Psacharopoulos (1994)
Chipre	1994	5,2	Menon (1995)
República Checa	1995	8,1	Filler <i>et al.</i> (1999)
República Checa	1996	8,5	Filler <i>et al.</i> (1999)
República Checa	1997	9,0	Filler <i>et al.</i> (1999)
Dinamarca	1976-1984	2,6	Christensen y Westergard-Nielsen (1999)
Dinamarca	1990	4,5	Christensen y Westergard-Nielsen (1999)
El Salvador	1985	7,9	Funkhouser (1996)
El Salvador	1988	7,8	Funkhouser (1996)
El Salvador	1990	7,6	Funkhouser (1996)
El Salvador	1992	7,6	Funkhouser (1996)
Estonia*	1989	1,5	Kroncke y Smith (1999)
Estonia*	1994	5,4	Kroncke y Smith (1999)
Finlandia	1980	9,1	Asplund (1999)
Finlandia	1987	7,0	Asplund (1999)
Finlandia	1989	8,2	Asplund (1999)
Finlandia	1991	8,8	Asplund (1999)
Finlandia	1993	8,2	Asplund (1999)
Alemania	1986	5,5	Ichino y Winter-Ebmer (1999)
Alemania	1988	7,7	Cohn y Addison (1998)
Alemania (Oriental)	1989	4,4	Munich <i>et al.</i> (1999)
Alemania (Oriental)	1991	4,1	Munich <i>et al.</i> (1999)
Alemania (Occidental)	1977	12,6	Patrinos (1995)
Alemania (Occidental)	1987	4,9	Patrinos (1995)
Ghana	1989	8,5	Glewwe (1996)
Ghana	1995	7,1	Jones (2001)
Grecia	1964	8,6	Patrinos (1995)
Grecia	1977	5,8	Patrinos (1995)
Grecia	1985	7,9	Cohn y Addison (1998)
Grecia	1987	2,7	Patrinos (1995)
Grecia	1993	7,6	Magoula y Psacharopoulos (1999)

Tabla A.4 (continuación)

## RENDIMIENTOS DE LA INVERSIÓN EN EDUCACIÓN COMPARABLES EN EL TIEMPO EN PAÍSES SELECCIONADOS

PAÍS	AÑO	TASA DE RENDIMIENTO (%)	FUENTE
Guatemala	1977	12,7	Funkhouser (1996)
Guatemala	1986	9,8	Funkhouser (1996)
Guatemala	1989	14,9	Psacharopoulos (1994)
Honduras	1986	12,5	Bedi (1997)
Honduras	1989	11,5	Funkhouser (1996)
Honduras	1990	10,4	Funkhouser (1996)
Honduras	1991	9,3	Funkhouser (1996)
Indonesia	1981	17,0	Psacharopoulos (1994)
Indonesia	1995	7,0	Duflo (2001)
Italia	1977	4,5	Brunello <i>et al.</i> (1999)
Italia	1983	5,5	Brunello <i>et al.</i> (1999)
Italia	1985	4,5	Brunello <i>et al.</i> (1999)
Italia	1986	4,6	Bruneilo <i>et al.</i> (1999)
Italia	1987	2,7	Brunello <i>et al.</i> (1999)
Japón	1975	6,5	Psacharopoulos (1994)
Japón	1978	4,4	Cohn y Addison (1998)
Japón	1988	13,2	Cohn y Addison (1998)
Kenia	1970	16,4	Psacharopoulos (1994)
Kenia	1986	16,0	Dabalen (1998)
Corea	1974	12,0	Ryoo <i>et al.</i> (1993)
Corea	1979	14,1	Ryoo <i>et al.</i> (1993)
Corea	1986	13,5	Ryoo <i>et al.</i> (1993)
México	1984	6,5	Patrinos (1995)
México	1989	7,5	Patrinos (1995)
México	1991	16,1	Psacharopoulos <i>et al.</i> (1996)
México	1992	7,6	Psacharopoulos <i>et al.</i> (1996)
Holanda	1962	11,0	Hartog <i>et al.</i> (1999)
Holanda	1965	12,2	Hartog <i>et al.</i> (1999)
Holanda	1972	11,3	Hartog <i>et al.</i> (1999)
Holanda	1979	10,9	Hartog <i>et al.</i> (1999)
Holanda	1982	7,0	Cohn y Addison (1998)
Holanda	1985	7,2	Hartog <i>et al.</i> (1999)
Holanda	1986	5,2	Cohn y Addison (1998)
Holanda	1988	5,7	Cohn y Addison (1998)
Holanda	1989	7,3	Hartog <i>et al.</i> (1999)

Tabla A.4 (continuación)

### RENDIMIENTOS DE LA INVERSIÓN EN EDUCACIÓN COMPARABLES EN EL TIEMPO EN PAÍSES SELECCIONADOS

PAÍS	AÑO	TASA DE RENDIMIENTO (%)	FUENTE
Holanda	1994	6,4	Hartog <i>et al.</i> (1999)
Nicaragua	1985	6,5	Funkhouser (1996)
Nicaragua	1993	7,9	Funkhouser (1996)
Nicaragua	1996	12,1	Belli y Ayadi (1998)
Noruega	1980	5,5	Barth y Roed (1999)
Noruega	1983	6,1	Barth y Roed (1999)
Noruega	1987	5,4	Barth y Roed (1999)
Noruega	1989	4,9	Barth y Roed (1999)
Noruega	1991	5,4	Barth y Roed (1999)
Noruega	1995	5,5	Barth y Roed (1999)
Pakistán	1986	4,6	Psacharopoulos (1994)
Pakistán	1991	15,4	Katsis <i>et al.</i> (1998)
Panamá	1983	12,1	Patrinos (1995)
Panamá	1989	13,7	Patrinos (1995)
Panamá	1990	13,7	Psacharopoulos (1994)
Paraguay	1983	11,6	Patrinos <i>et al.</i> (1994)
Paraguay	1990	11,5	Patrinos <i>et al.</i> (1994)
Perú	1985	11,5	Patrinos (1995)
Perú	1990	8,1	Psacharopoulos (1994)
Perú	1991	5,7	Patrinos (1995)
Filipinas	1982	8,0	Patrinos (1995)
Filipinas	1988	8,0	Patrinos (1995)
Filipinas	1994	7,3	Maluccio (1998)
Filipinas	1998*	12,6	Schady (2000)
Polonia	1986	2,9	Lorenz y Wagner (1993)
Polonia	1987	5,0	Rutkowski (1997)
Polonia	1992	7,0	Rutkowski (1997)
Polonia	1995-1996	7,0	Nesterova y Sabirianova (1998)
Portugal	1977	9,1	Patrinos (1995)
Portugal	1985	10,0	Patrinos (1995)
Portugal	1991	8,6	Cohn y Addison (1998)
Federación de Rusia	1995	7,5	Nesterova y Sabirianova (1998)
Federación de Rusia	1996	7,2	Nesterova y Sabirianova (1998)
Singapur	1974	13,4	Psacharopoulos (1994)
Singapur	1998	13,1	Sakellariou (2003)

Tabla A.4 (continuación)

## RENDIMIENTOS DE LA INVERSIÓN EN EDUCACIÓN COMPARABLES EN EL TIEMPO EN PAÍSES SELECCIONADOS

PAÍS	AÑO	TASA DE RENDIMIENTO (%)	FUENTE
Sudáfrica	1994	4,1	Dabalen(1998)
Corea del Sur	1974	12,0	Ryoo <i>et al.</i> (1993)
Corea del Sur	1976	6,5	Patrinos (1995)
Corea del Sur	1979	14,1	Ryoo <i>et al.</i> (1993)
Corea del Sur	1980	11,1	Patrinos (1995)
Corea del Sur	1986	13,5	Ryoo <i>et al.</i> (1993)
España	1985	7,7	Cohn y Addison (1998)
España	1990	9,0	Alba-Ramirez y Segundo (1995)
España	1991	7,2	Mora (1999)
Suecia	1968	7,5	Arai y Kjellstrom (1999)
Suecia	1974	4,3	Arai y Kjellstrom (1999)
Suecia	1981	3,5	Arai y Kjellstrom (1999)
Suecia	1984	3,9	Arai y Kjellstrom (1999)
Suecia	1990	4,5	Isacsson (1999)
Suiza	1987	7,9	Psacharopoulos (1994)
Suiza	1990	7,5	Weber y Wolter (1999)
Tailandia	1971	10,4	Psacharopoulos (1994)
Tailandia	1986	12,4	Patrinos (1995)
Tailandia	1989	11,5	Patrinos (1995)
Reino Unido	1975	8,0	Patrinos (1995)
Reino Unido	1982	15,3	Harmon y Walker (1995)
Reino Unido	1984	13,3	Harmon y Walker (1999)
Reino Unido	1987	6,5	Patrinos (1995)
Estados Unidos	1976	7,5	Kling (1999)
Estados Unidos	1978	7,9	Patrinos (1995)
Estados Unidos	1987	9,8	Psacharopoulos (1994)
Estados Unidos	1991-1995	10,0	Psacharopoulos (2000)
Venezuela	1975	13,7	Patrinos (1995)
Venezuela	1984	11,2	Patrinos (1995)
Venezuela	1989	9,6	Fiszbein y Psacharopoulos (1993)
Venezuela	1992	9,4	Psacharopoulos y Mattson (1998)
Vietnam	1992	4,8	Mooock et al, (1998)
Vietnam (Sur)	1964	16,8	Psacharopoulos (1994)
Yugoslavia	1976	6,8	Bevc (1993)
Yugoslavia	1986	4,8	Bevc (1993)

\* Datos para la población de hombres solamente.



3

## LA RENTABILIDAD DE LA EDUCACIÓN: ENFOQUE MICROECONÓMICO

Colm Harmon\*  
Hessel Oosterbeek\*\*  
Ian Walter\*\*\*

---

\* University College Dublin y CEPR.

\*\* Instituto Tinbergen, Universidad de Amsterdam.

\*\*\* Universidad de Warwick IFS.



### 3. LA RENTABILIDAD DE LA EDUCACIÓN: ENFOQUE MICROECONÓMICO\*

#### ■ RESUMEN

Este documento se centra en la educación como una decisión privada de inversión en “capital humano” y en la estimación de la tasa de rendimiento de esa inversión privada. A pesar de que la literatura está llena de estudios que estiman la tasa de rendimiento utilizando métodos de regresión donde los rendimientos estimados se han obtenido como el coeficiente de la variable años de educación en la ecuación del logaritmo del salario que incluye controles por experiencia laboral y otras características personales, tales aproximaciones no están exentas de dificultades. Se ofrecen argumentos teóricos que sustentan los desarrollos empíricos, y se demuestra que la evidencia sobre la rentabilidad individual privada es convincente. A pesar de algunos problemas con relación a la estimación de los rendimientos educativos, la evidencia, basada en estimaciones para una variedad de bases de datos y especificaciones, es que existe un claro efecto positivo derivado de la decisión individual de educarse sobre los ingresos. Adicionalmente, la magnitud del efecto es elevado frente a la rentabilidad de otras inversiones.

**Palabras clave.** Rentabilidad de la educación, endogeneidad, señalización, capital humano.

#### ■ 3.1. INTRODUCCIÓN

El objeto de interés de este artículo son los rendimientos de la educación. En particular nos centramos en la educación como decisión privada de invertir en “capital humano” y examinamos la tasa “interna” de rendimiento de esta inversión privada. Aunque en la literatura abundan estudios que estiman la tasa de rendimiento utilizando métodos de regresión, donde los rendimientos estimados se obtienen como el coeficiente de la variable años de educación en la ecuación del logaritmo del salario, que incluye controles por experien-

---

\* Colm Harmon, Hessel Oosterbeek e Ian Walker, “The Returns to Education: Microeconomics”, *Journal of Economics Surveys*, vol. 17, n.º 2, 2003, páginas 115-155.

cia laboral y otras características personales, el problema no está exento de dificultades. La evidencia de que los rendimientos privados son desproporcionadamente altos con relación a otras inversiones con un grado similar de riesgo, sugeriría que existe algún “fallo en el mercado” que impide que los individuos optimicen sus decisiones. Ello podría justificar ciertas intervenciones. Otro argumento para la intervención sería la existencia de externalidades asociadas con los logros educativos de los individuos, y esta es la esencia de la contribución de Sianesi y Van Reenen (2002).

En la siguiente sección se ofrecen los argumentos teóricos que sustentan los desarrollos empíricos, prestando particular atención a varias de las dificultades empíricas más importantes. En la sección tercera se revisan los trabajos existentes y se exploran estimaciones para diversas especificaciones y bases de datos del Reino Unido. La sección cuatro considera el papel de la señalización en la educación y el efecto del credencialismo en los salarios. Un problema relacionado, discutido en la sección quinta, es la magnitud de la heterogeneidad en los rendimientos educativos: los rendimientos pueden variar entre individuos ya que difiere la eficiencia con la que aprovechan la educación para elevar su productividad. Finalmente, la sección seis concluye.

## ■ 3.2. EL ENFOQUE DE CAPITAL HUMANO Y LA RENTABILIDAD DE LA EDUCACIÓN

### ■ 3.2.1. La especificación de Mincer

El análisis de la demanda de educación ha sido impulsado por el concepto de capital humano, introducido por Gary Becker, Jacob Mincer y Theodore Schultz. En la teoría del capital humano, la educación es una inversión de recursos actuales (los costos de oportunidad del tiempo involucrado, así como los costos directos), a cambio de rendimientos futuros. El modelo de referencia para el desarrollo de la estimación empírica de los rendimientos de la educación es la relación clave derivada por Mincer (1974). La teoría tradicional del capital humano (Becker, 1964) supone que la educación,  $s$ , es elegida para maximizar el valor presente esperado del flujo de ingresos futuros  $w$ , hasta la edad de jubilación  $T$ , neto de los costos de educación,  $c_3$ . Así, en el  $s$  óptimo, el valor presente del  $s$  —ésimo año de escolaridad es igual al costo del  $s$  —ésimo año de educación, y el equilibrio se caracteriza por:

$$\sum_{t=1}^{T-s} \frac{W_s - W_{s-1}}{(1+r_s)^t} = W_{s-1} + C_3$$

donde  $r_s$  se denomina la tasa interna de rentabilidad (suponemos que  $s$  es infinitamente divisible, por simplicidad, entonces “año” no debería ser interpretado literalmente). Tomar la decisión de inversión óptima implicaría que se invertiría en  $s$  años de escolaridad si  $r_s > i$ , siendo “ $i$ ” el tipo de interés del mercado. Si  $T$  es elevado entonces la relación de equilibrio del lado izquierdo puede aproximarse de forma que la condición de equilibrio se convierte en:

$$\frac{W_s - W_{s-1}}{r_s} = W_{s-1} + C_s$$

Entonces, si  $c_s$  es suficientemente pequeño, podemos reordenar esta expresión como:

$$r_s \approx \frac{W_s - W_{s-1}}{W_{s-1}} \approx \log w_s - \log w_{s-1}$$

(donde  $\approx$  significa aproximadamente igual a). Esto implica que el rendimiento del  $s$ -ésimo año de escolaridad es aproximadamente la diferencia en el logaritmo de los salarios entre  $s$  y  $s - 1$ . Entonces, se pueden estimar los rendimientos de  $s$  analizando como cambia el logaritmo del salario con  $s$ .

La contrapartida empírica del modelo teórico de capital humano es la familiar forma funcional de la ecuación de salarios.

$$\log w_i = X_i \beta + r s_i + \delta x_i + \gamma x_i^2 + u_i$$

donde  $w_i$  es una medida del ingreso para el individuo  $i$  dado como el ingreso por hora o por semana,  $s_i$  representa la medida de su escolaridad,  $x_i$  es la medida de experiencia,  $X_i$  es un conjunto de otras variables que afectan al ingreso, y  $u_i$  es el término de perturbación aleatoria que representa otros componentes que no han sido explícitamente incluidos y que se suponen independientes de  $X_i$  y  $s_i$ . Cabe destacar que la experiencia se introduce como una expresión cuadrática para captar la concavidad del perfil de ingresos frente a la experiencia. La derivación de Mincer del modelo empírico implica que, bajo los supuestos planteados (particularmente ausencia de costes de matrícula),  $r$  puede ser considerado como el rendimiento financiero privado de la educación, a la vez que capta el efecto sobre los salarios derivado de un incremento en  $s$ .

La disponibilidad de microdatos y las facilidades de estimación han dado lugar a muchos estudios, que estiman la sencilla especificación de Mincer. En el estudio original de Mincer (1974) su autor utilizó los datos del Censo de Estados Unidos de 1960 y una medida de la experiencia conocida como experiencia potencial (es decir la edad actual menos la edad a la que el individuo dejó de estudiar a tiempo completo), y encontró que la rentabilidad de la educación era de un 10%, con una rentabilidad de la experiencia alrededor del 8%. Psacharopoulos y Layard (1979) usaron los datos del *GB GHS* de 1972, y de igual forma encontraron rendimientos educativos alrededor del 10%. Véase Willis (1986) y Psacharopoulos (1994) para más ejemplos de la especificación simple!

### ■ 3.2.2. Elección óptima de escolaridad

En el trabajo empírico discutido anteriormente, la escolaridad se trata como exógena, a pesar de que de acuerdo con la teoría del capital humano el nivel educativo selecciona-

do es una variable endógenamente determinada. En este sentido, es útil considerar las implicaciones de la endogeneidad de la educación. Como se sugirió antes, en el enfoque del capital humano sobre el que se basa el trabajo original de Mincer, la escolaridad es una decisión de inversión basada en la optimización de las ganancias futuras y los costes corrientes: es decir, sobre el valor presente de la diferencia de ingresos entre las alternativas de educarse o no y el coste total de la educación incluyendo los ingresos no percibidos. La inversión en educación continúa hasta que la diferencia entre el costo marginal y el rendimiento marginal de la educación se hace cero.

La consideración de la escolaridad como una decisión tiene un conjunto de implicaciones. Primero, la tasa interna de rentabilidad (TIR, o  $r$  en este enfoque) es la tasa de descuento que iguala el valor presente de los beneficios y el valor presente de los costos. Más concretamente si la TIR es mayor que el tipo de interés del mercado (suponiendo que un individuo puede pedir prestado a este tipo), obtener educación adicional es una inversión rentable para el individuo. Al tomar una decisión de inversión, un individuo que concede más (menos) valor al ingreso corriente que a los flujos de ingreso futuro tendrá un mayor (menor) valor para la tasa de descuento. En consecuencia, individuos con altas tasas de descuento (alto  $r_i$ ) estarán *menos* dispuestos a adquirir educación<sup>11</sup>. Segundo, los costes directos de la educación ( $c_s$ ) reducen el beneficio neto de educarse. Tercero, si la probabilidad de estar empleado es más alta a una mayor escolarización, un incremento en el subsidio por desempleo podría reducir los incentivos a educarse. No obstante, caso de que la brecha de ingresos entre individuos educados y no educados se ampliase o si los costes de oportunidad de educarse se redujesen (por ejemplo, a través de subsidios o de ayudas), el efecto neto sobre el incentivo a invertir en escolaridad debería ser positivo. Cuarto, más escolaridad puede implicar una mayor probabilidad de recibir formación en el puesto de trabajo en la medida en que ambos tipos de formación sean complementarios (Blundell *et al.*, 1996). Quinto, pueden existir beneficios no pecuniarios asociados con la educación, que no se reflejan en los salarios, incluyendo los asociados con tener un trabajo más cualificado (Chevalier y Lydon, 2001). Finalmente, Heckman *et al.* (1999) planteó la diferencia entre análisis de equilibrio parcial y de equilibrio general, donde en el último caso la distribución del salario bruto cambia en el sentido en que parcialmente compensa el efecto de algún cambio de política, a través de la incidencia sobre el lado de la demanda en el mercado. Por lo tanto, a menos que la demanda laboral sea perfectamente elástica para todo tipo de empleo, los aumentos en los incentivos individuales a invertir en educación, dada la distribución actual del salario, se verán compensados por cambios en la distribución cuando la oferta laboral de educados aumente y la oferta de no educados disminuya.

Una extensión útil de la teoría es considerar la contribución de la habilidad individual sobre la decisión de educarse, preservando al mismo tiempo la idea básica de la educación como una inversión. Griliches (1977) introduce la habilidad ( $A$ ) explícitamente en la derivación de la función de ingresos log-lineal. En el modelo básico, la TIR de la educación se determina parcialmente por los ingresos no percibidos (menos los subsidios gubernamen-

tales o la contribución de los padres) y todos los costos educativos. Introducir diferencias en habilidad tiene dos efectos en el cálculo básico. Los individuos más hábiles puede que sean capaces de convertir educación en capital humano de forma más eficiente que los menos hábiles, y ello se traduce en un aumento de la TIR para los primeros<sup>III</sup>. Cabe pensar en la habilidad individual innata y la educación como factores complementarios en la producción de capital humano, lo que se traduce en que, para un aumento dado de escolaridad, una mayor dotación de habilidad innata genere un mayor aumento de capital humano. Por otro lado, los individuos más hábiles pueden tener un coste de oportunidad más elevado dado que son capaces de obtener en el mercado una retribución más elevada, en la medida en que la capacidad para progresar académicamente esté positivamente correlacionada con la capacidad para obtener salarios más altos, y ello debe contribuir a reducir la TIR<sup>IV</sup>.

Las consecuencias empíricas de esta extensión de la teoría básica, se exponen con mayor claridad por Card (1999), que a su vez introduce la idea de que el nivel de escolaridad óptimo equilibra la tasa de rendimiento marginal de la educación adicional, con el coste marginal de esta escolaridad. No obstante, Card (1999) permite que la escolaridad óptima varíe entre individuos por una razón adicional: no sólo pueden aparecer diferencias en los rendimientos educativos a consecuencia de la habilidad en el sentido de que los más hábiles obtienen un mayor beneficio de la escolaridad extra, sino que también aparecen diferencias en las tasas marginales de sustitución entre ingresos corrientes y futuros. Es decir, puede haber alguna variación en la tasa de descuento entre individuos. Esta variación en la tasa de descuento puede venir explicada, por ejemplo, por las diferencias en la facilidad de acceso a los recursos o por las preferencias por la educación (Lang, 1993).

Si el nivel de habilidad es similar entre individuos, entonces el efecto es relativamente directo —a menor tasa de descuento los individuos escogen mayor escolaridad. No obstante, cabría esperar que una correlación negativa entre estos dos elementos: unos padres muy hábiles, que generalmente serán también más ricos, serán capaces de ofrecer a sus hijos más recursos para educarse. Adicionalmente, padres altamente educados tendrán un mayor gusto por la educación (o más baja tasa de descuento) y sus hijos “heredarán” parte de estas preferencias. De hecho, si la habilidad es parcialmente heredada, entonces los niños con mayor habilidad pueden ser más propensos a tener tasas de descuento más bajas que los niños promedio. Lo contrario es cierto para niños con padres menos hábiles. Empíricamente, esta modificación permite que pueda derivarse una expresión del sesgo potencial en la estimación mínimo cuadrática de los rendimientos educativos. Este sesgo estará determinado por la varianza en la habilidad respecto a la varianza en la tasa de descuento, así como por la covarianza entre ellas. Este sesgo de “endogeneidad” surge porque individuos con mayor rendimiento marginal (o menor coste marginal) de la educación escogerán mayores niveles de escolaridad. Si no hay varianza en la tasa de descuento entonces la endogeneidad surgirá sólo a consecuencia de la correlación entre habilidad y educación, y dado que esta tiende a ser positiva, el sesgo en la estimación MCO será al alza (si la habilidad incrementa los salarios al final de la vida más de lo que los incrementa al principio de

la vida). Si no hay varianza en la habilidad, entonces la endogeneidad surgirá solamente de la correlación (negativa) entre tasas de descuento y la cantidad de educación, y la estimación MCO será sesgada a la baja si las tasas de descuento y los salarios están positivamente correlacionados (por ejemplo, si personas ambiciosas ganan un mayor salario y son más impacientes). Por lo tanto, el sentido del sesgo en la estimación MCO de los rendimientos educativos no está claro, y es, en última instancia, una cuestión empírica.

### ■ 3.2.3. Sesgo de habilidad

En la especificación de Mincer la perturbación aleatoria recoge efectos individuales inobservables, y estos efectos individuales pueden además influir sobre la decisión de educarse, y por lo tanto, introducir una correlación entre la escolaridad y el término de error en la función de ingresos. Un ejemplo común es la habilidad inobservada. Este problema ha constituido una preocupación en la literatura empírica desde las primeras contribuciones —si la escolaridad es endógena entonces la estimación por el método de mínimos cuadrados ofrecería estimaciones sesgadas del rendimiento de la educación.

Existe una serie de enfoques para tratar este problema. Primero, se han incorporado medidas de habilidad como proxy del efecto inobservado. La inclusión de medidas directas de habilidad debería reducir el coeficiente de la educación estimado, si actúan como proxy de la habilidad, de modo que el coeficiente de la educación recoja únicamente el efecto de la educación dado que la habilidad se mantiene<sup>v</sup>. Segundo, cabe explotar las diferencias en salarios y educación entre gemelos (o entre hermanos), si se acepta la hipótesis de que los efectos inobservados son aditivos y comunes entre gemelos, de modo que pueden ser eliminados al calcular una regresión de las diferencias de salarios entre gemelos frente a las diferencias en educación. Un último enfoque considera directamente la estimación de un sistema de dos ecuaciones, una para ingresos y otra para escolaridad. La identificación del sistema se logra sobre la base de aquellas variables que afectan directamente a la escolaridad pero no a los salarios. Volvemos con mayor detalle sobre esta cuestión al final del artículo.

## ■ 3.3. ANÁLISIS DE REGRESIÓN

### ■ 3.3.1. Análisis comparativo

Dado que los salarios se determinan por una serie de variables, algunas de las cuales están correlacionadas entre sí, además de con los salarios, necesitamos usar métodos de regresión multivariante para derivar estimaciones válidas del efecto sobre los salarios de cualquier variable —en particular, de la educación. La tabla 3.1 presenta las estimaciones de la tasa de rendimiento de la educación basado en el análisis multivariante MCO con los datos del *International Social Survey Programme* (ISSP), que se realizan a partir de las

Tabla 3.1

### EVIDENCIA DE LOS RENDIMIENTOS EDUCATIVOS ENTRE PAÍSES-ISSP 1995

PAÍS	HOMBRES		MUJERES	
Australia	0,0509	<i>0,0042</i>	0,0568	<i>0,0077</i>
Alemania Occidental	0,0353	<i>0,0020</i>	0,0441	<i>0,0036</i>
Gran Bretaña	0,1299	<i>0,0057</i>	0,1466	<i>0,0069</i>
Estados Unidos	0,0783	<i>0,0045</i>	0,0979	<i>0,0058</i>
Austria	0,0364	<i>0,0033</i>	0,0621	<i>0,0049</i>
Italia	0,0398	<i>0,0025</i>	0,0568	<i>0,0036</i>
Hungría	0,0699	<i>0,0053</i>	0,0716	<i>0,0051</i>
Suiza	0,0427	<i>0,0065</i>	0,0523	<i>0,0143</i>
Polonia	0,0737	<i>0,0044</i>	0,1025	<i>0,0046</i>
Holanda	0,0331	<i>0,0025</i>	0,0181	<i>0,0050</i>
República de Irlanda	0,1023	<i>0,0051</i>	0,1164	<i>0,0081</i>
Israel	0,0603	<i>0,0069</i>	0,0694	<i>0,0077</i>
Noruega	0,0229	<i>0,0025</i>	0,0265	<i>0,0032</i>
Irlanda del Norte	0,1766	<i>0,0111</i>	0,1681	<i>0,0127</i>
Alemania Oriental	0,0265	<i>0,0032</i>	0,0450	<i>0,0041</i>
Nueva Zelanda	0,0424	<i>0,0050</i>	0,0375	<i>0,0058</i>
Rusia	0,0421	<i>0,0042</i>	0,0555	<i>0,0043</i>
Eslovenia	0,0892	<i>0,0104</i>	0,1121	<i>0,0091</i>
Suecia	0,0367	<i>0,0047</i>	0,0416	<i>0,0047</i>
Bulgaria	0,0495	<i>0,0100</i>	0,0624	<i>0,0091</i>
Canadá	0,0367	<i>0,0072</i>	0,0498	<i>0,0083</i>
República Checa	0,0291	<i>0,0069</i>	0,0454	<i>0,0077</i>
Japón	0,0746	<i>0,0066</i>	0,0917	<i>0,0151</i>
España	0,0518	<i>0,0071</i>	0,0468	<i>0,0099</i>
Eslovaquia	0,0496	<i>0,0070</i>	0,0635	<i>0,0078</i>

Nota: Errores estándar en *itálica*.

encuestas individuales pero con un diseño conjunto que garantiza su comparabilidad. Por ejemplo, los datos británicos del ISSP se toman de la *British Social Attitudes Survey*. En la tabla 3.1, se aplica exactamente el mismo método de estimación para facilitar la comparabilidad entre países. Estos resultados muestran importantes diferencias entre países.

Estas estimaciones tienen la ventaja de haber sido obtenidas a partir de datos comunes lo que, en términos generales, las hace comparables si bien incurren en el coste de la simplicidad. En particular, los modelos estimados contienen controles sólo por edad, y por si el individuo está o no sindicado —la inclusión de variables de control adicionales, proba-

blemente tenderían a reducir el coeficiente estimado para la escolaridad. Adicionalmente, los datos del ISSP han sido diseñados para el análisis cualitativo, y parece probable por tanto, que pueda haber errores de medida en los ingresos o en la escolaridad. A consecuencia del error de medida, en general, la estimación de los rendimientos educativos estará sesgada a la baja, por lo que hay que tener cautela en la interpretación de estos resultados<sup>VI</sup>. En este sentido, podría ser interesante considerar las tasas de rendimiento entre países obtenidas de las encuestas nacionales más que de una sola fuente homogénea como la ISSP. Resultados recientes de la red de investigadores Paneuropeos (conocida con las siglas de PURE “*Public Funding and Private Returns to Education*”) siguen esta estrategia y obtienen las estimaciones a partir de las fuentes de datos de cada país, lo que permite aprovechar los mejores datos. El objetivo principal era evaluar el rendimiento privado de la educación estimando la relación entre salarios y educación en Europa. En un proyecto entre países es preferible que los datos sean razonablemente comparables entre países, es decir, los salarios, los años de escolaridad y la experiencia deben haber sido calculados de forma similar. Sin embargo, dado que cada país utiliza sus propias fuentes de datos, esta condición es difícil de que se satisfaga plenamente. Todos los miembros del proyecto PURE adoptaron una especificación común, y estimaron los rendimientos educativos usando el logaritmo del salario bruto cuando se disponía de la variable<sup>VII</sup>.

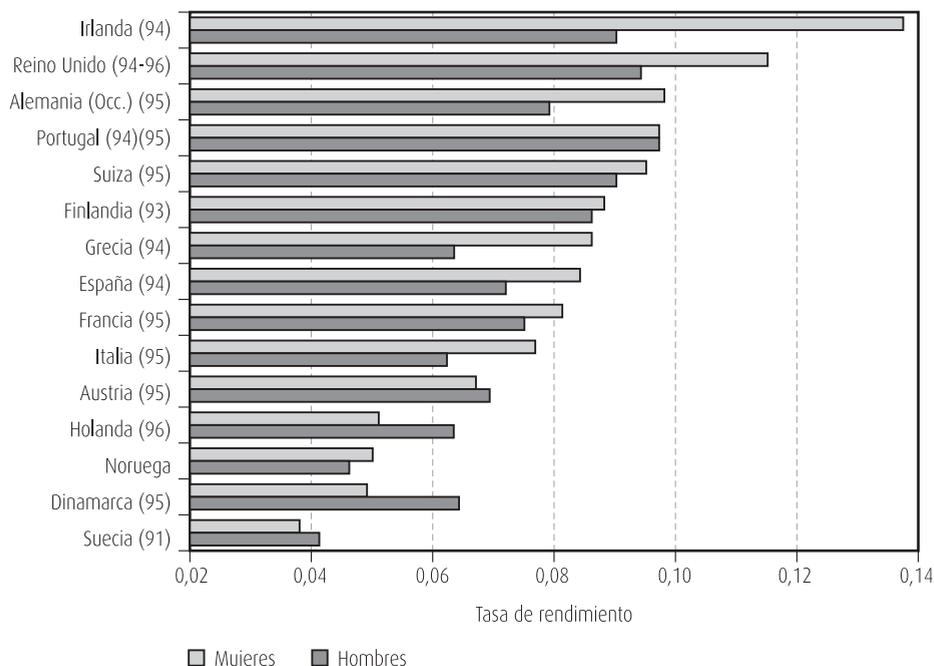
El gráfico 3.1 es un resumen de los rendimientos por género. Fueron obtenidos a partir de una especificación simple que contenía sólo la escolaridad y un término cuadrático en la edad. Encontramos que para algunos países como el Reino Unido, Irlanda, Alemania, Grecia e Italia, existen diferencias sustanciales en los rendimientos entre sexos, —los rendimientos para mujeres son significativamente más altos que los rendimientos para hombres. Escandinavia (Noruega, Suecia y Dinamarca) se caracterizan por tener rendimientos relativamente bajos. Irlanda y el Reino Unido están cerca del límite superior de los rendimientos estimados en esta comparación entre países.

### ■ 3.3.2. Especificación y forma funcional

La especificación de Mincer puede ser considerada como una aproximación a la función más general de la escolaridad ( $s$ ) y la experiencia ( $x$ ) que adopta la forma:  $\log w = F(s, x) + e$ , donde  $e$  es la perturbación aleatoria que capta otros determinantes (inobservables) de los salarios. Se han probado muchos variantes para la especificación de  $F(\cdot)$ . Murphy y Welch (1990), por ejemplo, concluyen que el  $\log w = X\beta + rs + g(x) + e$ , donde  $X$  son las características observables que afectan los salarios y  $g(\cdot)$  es un polinomio de tercer o cuarto orden de la experiencia, ofrece la mejor aproximación. No obstante, existen pocos ejemplos en la literatura empírica, que consideren si la forma en que  $x$  entra en el modelo tiene un impacto sustancial sobre el coeficiente de escolaridad estimado. Kjellstrom y Björklund (2001) no observan efectos empleando datos suecos, mientras Heckman, Lochner y Todd (2001) empleando datos del Censo de EE.UU. sugieren que la no considera-

Gráfico 3.1

### RENDIMIENTOS EDUCATIVOS EN EUROPA, HOMBRES Y MUJERES (AÑO CERCANO A 1995)



Fuente: Trostel, Walker y Wooley (2002). La especificación de la regresión incluye controles por edad, edad al cuadrado y unión sindical.

ción de interacciones entre la experiencia y la escolaridad tiene importantes implicaciones sobre la estimación de la tasa de rendimiento de la escolaridad, al menos para los últimos datos del censo.

Sin embargo, la experiencia rara vez aparece medida en las bases de datos típicas y es frecuente su aproximación a partir de la diferencia entre la edad y la edad en que dejó de estudiar, o incluso simplemente por la edad. Es preciso tener en cuenta que para comparar la especificación que usa la edad con las que usan la experiencia observada o la experiencia potencial se necesita ajustar por la diferencia en lo que en las distintas especificaciones se supone que se mantiene constante. El efecto de  $s$  sobre el logaritmo del salario —manteniendo la experiencia constante es simplemente  $r$ , mientras que la especificación que controla por la edad implica que la estimación del impacto que tiene la educación sobre los salarios manteniendo la edad constante necesita ser ajustada por el efecto de  $s$  sobre la experiencia— es decir, se necesita restar el efecto de los años de experiencia<sup>viii</sup>.

La tabla 3.2 muestra el efecto de incluir diferentes medidas de experiencia en la estimación de los rendimientos educativos. En esta tabla presentamos las estimaciones MCO controlando por diferentes definiciones de experiencia, empleando nuestras estimaciones de los rendimientos educativos para Europa. Especificar una función cuadrática en la edad tiende a generar rendimientos inferiores. Usar la experiencia potencial (edad menos la edad a la que dejó de estudiar) o la experiencia real (generalmente calculada como la suma ponderada del número de años de trabajo a tiempo parcial y a tiempo completo, desde que dejó de estudiar a tiempo completo) indica un rendimiento de la educación ligeramente más elevado. Por ejemplo, las estimaciones para el Reino Unido empleando los datos del *Family Resources Survey* (FRS) son del 10% para hombres y del 12% para mujeres, comparado con el 8% y el 11% respectivamente, cuando se usa la edad como *proxy* de la experiencia. Sin embargo, el tamaño muestral es grande y las estimaciones son muy precisas, por lo que incluso estas pequeñas diferencias son generalmente estadísticamente significativas<sup>IX</sup>.

Otros cambios en la especificación, generalmente no conllevan a grandes diferencias en los rendimientos educativos estimados. Por ejemplo, en las tablas 3.3 y 3.4 se estima-

Tabla 3.2

### RENDIMIENTOS EDUCATIVOS EN EUROPA (AÑO MÁS CERCANO A 1995)

DEFINICIÓN DE CONTROL POR EXPERIENCIA	HOMBRES			MUJERES		
	EXPERIENCIA POTENCIAL	EXPERIENCIA REAL	EDAD	EXPERIENCIA POTENCIAL	EXPERIENCIA REAL	EDAD
Austria (95)	0,069		0,059	0,067		0,058
Dinamarca (95)	0,064	0,061	0,056	0,049	0,043	0,044
Alemania (Occid) (95)	0,079	0,077	0,067	0,098	0,095	0,087
Holanda (96)	0,063	0,057	0,045	0,051	0,042	0,037
Portugal (94)(95)	0,097	0,100	0,079	0,097	0,104	0,077
Suecia (91)	0,041	0,041	0,033	0,038	0,037	0,033
Francia (95)	0,075		0,057	0,081		0,065
Reino Unido (94-96)	0,094	0,096	0,079	0,115	0,122	0,108
Irlanda (94)	0,090	0,088	0,065	0,137	0,129	0,113
Italia (95)	0,062	0,058	0,046	0,077	0,070	0,061
Noruega	0,046	0,045	0,037	0,050	0,047	0,044
Finlandia (93)	0,086	0,085	0,072	0,088	0,087	0,082
España (94)	0,072	0,069	0,055	0,084	0,079	0,063
Suiza (95)	0,090	0,089	0,076	0,095	0,089	0,086
Grecia (94)	0,063		0,040	0,086		0,064
<b>Media</b>	<b>0,073</b>	<b>0,072</b>	<b>0,058</b>	<b>0,081</b>	<b>0,079</b>	<b>0,068</b>

Fuente: Información recolectada en el grupo PuRE por Rita Asplund (ETLA, Helsinki).

Tabla 3.3

### SENSIBILIDAD A LOS CAMBIOS EN LAS VARIABLES DE CONTROL: HOMBRES EN LA BHPS

CONTOLES	NINGUNA	TAMAÑO EMPRESA, SINDICATO	HIJOS Y ESTADO CIVIL	TIEMPO PARCIAL	HIJOS, EST. CIVIL Y T. PARCIAL	TAMAÑO EMPRESA, SINDICATO Y T. PARCIAL	TODOS LOS CONTOLES
Educación	0,064 (0,002)	0,062 (0,002)	0,065 (0,002)	0,064 (0,002)	0,065 (0,002)	0,062 (0,002)	0,063 (0,002)
Empresa mediana	—	0,157 (0,012)	—	—	—	0,157 (0,012)	0,153 (0,012)
Empresa grande	—	0,241 (0,013)	—	—	—	0,242 (0,012)	0,243 (0,013)
Miembro sindicato	—	0,079 (0,011)	—	—	—	0,079 (0,011)	0,080 (0,011)
N.º de hijos	—	—	0,017 (0,006)	—	0,017 (0,006)	—	0,019 (0,005)
Casado	—	—	0,144 (0,016)	—	0,145 (0,016)	—	0,144 (0,016)
Pareja de hecho	—	—	0,095 (0,020)	—	0,095 (0,020)	—	0,107 (0,020)
Divorciado	—	—	0,050 (0,025)	—	0,050 (0,025)	—	0,058 (0,024)
Tiempo parcial	—	—	—	-0,020 (0,041)	-0,007 (0,041)	0,024 (0,039)	0,036 (0,040)

Fuente: Entre paréntesis los errores estándar robustos. El modelo incluye edad, edad al cuadrado, *dummies* de año, *dummies* de región y tasas de desempleo regional.

ron los rendimientos para hombres y mujeres empleando información del *British Household Panel Survey* (BHPS), incluyendo una serie de controles diferentes como pertenencia a un sindicato, tamaño de la empresa, trabajador a tiempo parcial, estado civil y tamaño del hogar<sup>x</sup>. Como puede observarse estos resultados son bastante robustos a cambios en la especificación.

Otro punto se refiere a la cuestión de emplear muestras de trabajadores ocupados con el fin de estimar los rendimientos educativos. ¿En qué medida la utilización de una muestra de ocupados produce sesgo en la estimación de los rendimientos educativos? Generalmente se considera que este no es un problema tan importante para los hombres, como lo es para las mujeres, dado que la no participación en el mercado laboral es mucho menor para hombres que para mujeres. Sin embargo, este argumento es menos realista para las cohortes recientes. Una forma sencilla de abordar el problema, puede ser usar el método

Tabla 3.4

### SENSIBILIDAD A LOS CAMBIOS EN LAS VARIABLES DE CONTROL: MUJERES EN LA BHPS

CONTOLES	NINGUNA	TAMAÑO EMPRESA, SINDICATO	HIJOS Y ESTADO CIVIL	TIEMPO PARCIAL	HIJOS, EST. CIVIL Y T. PARCIAL	TAMAÑO EMPRESA, SINDICATO Y T. PARCIAL	TODOS LOS CONTOLES
Educación	0,103 (0,002)	0,095 (0,002)	0,101 (0,002)	0,097 (0,002)	0,097 (0,002)	0,092 (0,002)	0,092 (0,002)
Empresa mediana		0,158 (0,010)	—	—	—	0,130 (0,010)	0,130 (0,010)
Empresa grande		0,258 (0,012)	—	—	—	0,217 (0,012)	0,216 (0,012)
Miembro sindicato		0,214 (0,012)	—	—	—	0,197 (0,012)	0,195 (0,012)
N.º de hijos	—		-0,077 (0,006)		-0,037 (0,006)	—	-0,032 (0,006)
Casado	—	—	0,001 (0,018)		0,029 (0,018)	—	0,025 (0,018)
Pareja de hecho	—	—	0,021 (0,022)		0,024 (0,022)	—	0,025 (0,021)
Divorciado	—	—	-0,009 (0,023)		-0,002 (0,022)	—	0,003 (0,021)
Tiempo parcial	—	—	—	-0,220 (0,009)	-0,197 (0,011)	-0,165 (0,009)	-0,156 (0,010)

Fuente: Entre paréntesis los errores estándar robustos. El modelo incluye edad, edad al cuadrado, *dummies* de año, *dummies* de región y tasas de desempleo regional.

de estimación estándar “en dos etapas” propuesto por Heckman y Polachek (1974), que pretende controlar por selección, modelando sus determinantes. La tabla 3.5 muestra la estimación del parámetro para mujeres empleando los datos Británicos de la BHPS y la FRS. Los resultados sugieren un efecto pequeño debido al sesgo de selección. Mientras la selección es estadísticamente significativa, las diferencias en los rendimientos educativos son pequeñas en valor absoluto, y no significativas.

Dado que la no participación es más común entre mujeres que entre hombres, podríamos suponer que los rendimientos de las mujeres estarían sesgados a la baja respecto al de los hombres, y la magnitud del sesgo podría depender de las tasas de participación relativas. El gráfico 3.2 examina la relación entre la tasa de participación promedio en el empleo para mujeres y el porcentaje de diferencia en los rendimientos educativos entre hombres y mujeres para los países de la red PURE. El gráfico muestra que los países con más alta tasa

Tabla 3.5

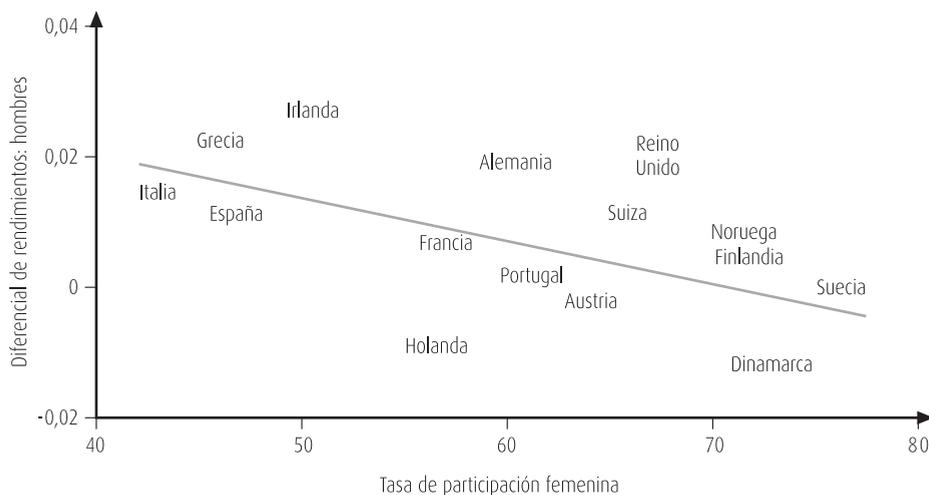
### ESTIMACIONES MCO Y SESGO DE SELECCIÓN DE HECKMAN: REINO UNIDO- BHPS Y FRS

MODELO	FRS MUJERES			BHPS MUJERES		
	EDUCACIÓN	EDAD	EDAD <sup>2</sup>	EDUCACIÓN	EDAD	EDAD <sup>2</sup>
MCO	<b>0,109</b> <b>(0,002)</b>	0,026 (0,003)	-0,0003 (0,00004)	<b>0,103</b> <b>(0,002)</b>	0,040 (0,005)	-0,0005 (0,0001)
Heckman	<b>0,109</b> <b>(0,002)</b>	0,016 (0,004)	-0,0001 (0,0001)	<b>0,102</b> <b>(0,003)</b>	0,060 (0,006)	-0,0007 (0,0001)

*Nota:* Entre paréntesis los errores estándar robustos. Los modelos incluyen *dummies* de año, estado civil, y el número de hijos en tres rangos de edad, *dummies* regionales, y tasa de desempleo regional. En el modelo Heckman en dos etapas se emplearon los ingresos no laborales del hogar, así como las variables de la educación de salarios en la ecuación de participación.

Gráfico 3.2

### DIFERENCIAS ENTRE RENDIMIENTOS DE HOMBRES Y MUJERES, Y TASA DE PARTICIPACIÓN FEMENINA



*Fuente:* Harmon, Walker y Westergaard-Nielsen (2001).

de participación femenina (generalmente el grupo nórdico) tienen las menores diferencias en rendimientos educativos, mientras que los países con menor participación (usualmente las economías mediterráneas) muestran las mayores diferencias. Irlanda y el Reino Unido (y en menor medida Alemania) presentan un comportamiento atípico, en el sentido en que aparecen relativamente elevadas diferencias en los rendimientos por género, mientras que están en el rango medio en términos de participación. No obstante, desde la perspectiva del investigador, esto puede sugerir algún sesgo potencial al usar muestras sólo de partici-

pantes, a pesar de que no parece ser un problema importante. Sin embargo, excepto para países con altas tasas de participación femeninas, el problema merece mayor atención del que ha recibido por parte de la literatura hasta la fecha.

### ■ 3.3.5. Medidas alternativas de la educación alcanzada

Medir la escolaridad en términos de años de educación cuenta con una larga tradición en los Estados Unidos. Existen razones prácticas como que los años de educación es una variable que figura en las principales bases de datos como el Censo y con anterioridad a 1990, en el *Current Population Survey* (CPS). Por otra parte, la educación en Estados Unidos no sigue un sistema nacional (o por estados) basado en credenciales sino que es un sistema donde los grados se definen por años, de modo que la educación es una variable bastante continua, al menos hasta que se llega a la graduación en secundaria. No obstante, en Europa existen vías alternativas de educación que pueden dar lugar a resultados muy diferentes. La estimación basada en títulos más que en años de escolaridad es, por lo tanto, una vía alternativa para obtener los rendimientos educativos. No obstante, esto es solo necesario si la respuesta salarial a incrementos en la educación se desvía de la linealidad en los años de educación. Consideremos una comparación de dos medidas de los rendimientos educativos, una basada en los años de escolaridad y otra basada en variables ficticias de niveles para el nivel más alto de escolaridad terminado. Si el rendimiento adicional (marginal) de un grado de tres años comparado con dejar la escuela con *A-level*<sup>1</sup>, es aproximadamente tres veces el rendimiento estimado de un año de *A-level*, entonces la especificación lineal en los años de escolaridad es equivalente a la alternativa basada en la titulación alcanzada.

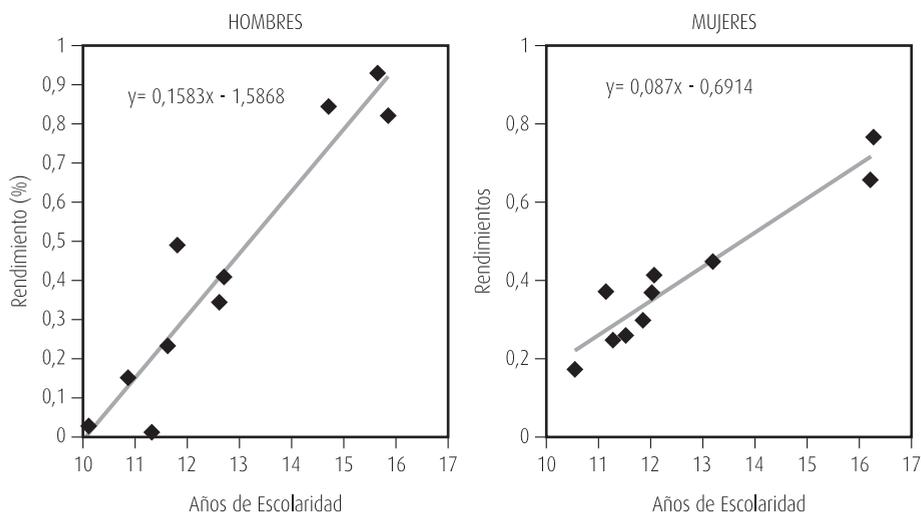
Algunos autores argumentan que los títulos importan más que los años de escolaridad —el llamado efecto “credencialismo” (“sheepskin effect”). Por ejemplo, puede existir una prima salarial sobre el rendimiento educativo promedio por la finalización de un año específico del currículum educativo (como, por ejemplo, el último año de la universidad o de secundaria). Hungerford y Solon (1987) demuestran la existencia de estas no linealidades. Park (1996) también observa una desviación de la linealidad en los rendimientos con relación a los años de escolaridad entre la finalización de la secundaria y la finalización de la universidad. Sus estimaciones sugieren que el rendimiento marginal de la educación no es constante sino que experimenta un cambio de pendiente entre estos dos puntos de transición importantes.

El gráfico 3.3 muestra que la hipótesis subyacente de linealidad, a pesar de ser restrictiva, resulta sorprendentemente difícil de rechazar. En este gráfico se ofrece el rendimiento promedio de un conjunto de títulos frecuentes en el Reino Unido (incluido aprendices, los grados profesionales nacionales y otras formas de educación) frente al número de años de escolaridad promedio para los poseedores de estos títulos. En caso de ajustar una

<sup>1</sup> (N. del T.) *A-level* representa un grado de dos años de duración.

Gráfico 3.3

## RENDIMIENTOS ESTIMADOS DE LAS CUALIFICACIONES-BHPS



regresión simple entre estos puntos observamos que la forma lineal parece ser una aproximación razonable, y que los rendimientos promedio para los años de escolaridad son alrededor de 16% para mujeres y 9% para hombres<sup>XI, XII</sup>.

### ■ 3.3.6. Variación de los rendimientos educativos a través de la distribución salarial

Es posible que los rendimientos educativos puedan diferir entre individuos, en la parte superior de la distribución salarial. Una de las propiedades de la estimación MCO es que la línea de regresión pasa a través de la media muestral. Una metodología alternativa disponible a la estimación MCO es conocida como Regresión Cuantílica (RC), que a pesar de basarse en toda la muestra, permite estimar los rendimientos educativos entre diferentes cuantiles de la distribución salarial (Buchinsky, 1994), al ponderar las observaciones de manera adecuada. Mientras los MCO captan el efecto de la educación sobre la media salarial, la idea que subyace a la RC es evaluar los rendimientos en alguna otra parte de la distribución salarial, por ejemplo, el cuartil inferior. Comparando las estimaciones de los rendimientos en el conjunto de la distribución salarial se puede inferir en qué medida la educación acentúa o reduce la desigualdad subyacente.

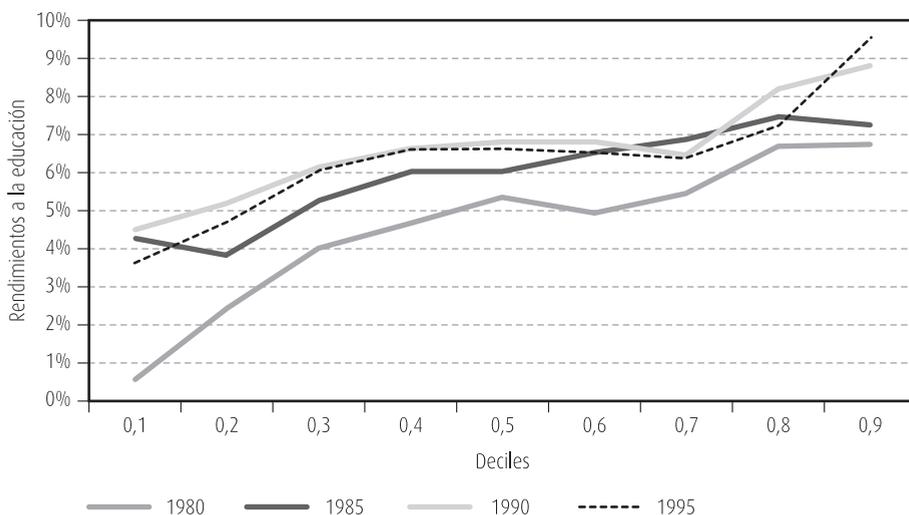
Por supuesto, el método requiere que la educación presente la suficiente variabilidad para poder identificar los rendimientos de cada decila se requiere que algunos en los deciles

superiores de ingresos tengan bajo nivel educativo, y que algunos deciles inferiores tengan alto nivel educativo. Los datos del Reino Unido parecen ser satisfactorios en este sentido, y se encontraron rendimientos estadísticamente significativos para cada decil, y además se encontró que los rendimientos para el decil superior eran significativamente más altos que para el decil inferior. El método es plenamente flexible y permite que los rendimientos para cada decil sean independientes de cualquier otro decil. La especificación exige que los rendimientos sean los mismos para todos los individuos que pertenecen a un decil, mientras que al aplicar MCO se supone que los rendimientos son los mismos para todos los individuos muestrales.

El gráfico 3.4 presenta la estimación MCO para el rendimiento promedio de la educación (empleando los datos del *Family Expenditure Survey* (FES) para los años 1980, 1985, 1990 y 1995), junto con los rendimientos educativos para diferentes deciles de la distribución salarial. Los resultados MCO muestran que en cuatro décadas los rendimientos educativos, en promedio, han aumentado en términos generales, especialmente entre 1980 y 1985. Hay una conclusión clara que se deduce de este gráfico, y es que los rendimientos educativos son mayores para aquellos en la parte superior de la distribución salarial en términos comparativos con los que se ubican en la parte inferior (a pesar de que los perfiles son planos para los rangos intermedios de la distribución). Aunque las diferencias no son importantes, los rendimientos de la parte inferior de la distribución parecen haber aumentado en el periodo, como se observa en la gráfica, disminuyendo la pendiente<sup>xiii</sup>. Comparando las décadas de los ochenta con los noventa también hay algún indicio de que los ren-

Gráfico 3.4

## REGRESIONES CUANTÍLICAS PARA GB: HOMBRES FES



Fuente: Harmon, Walker y Westergaard-Nielsen (2001).

Tabla 3.6

## REGRESIONES CUANTÍLICAS

PAÍS	AÑO	1º DECIL	9º DECIL	AÑO	1º DECIL	9º DECIL
Austria	1981	9,2	12,6	1993	7,2	12,8
Dinamarca	1980	4,7	5,3	1995	6,3	7,1
Finlandia	1987	7,3	10,3	1993	6,8	10,1
Francia	1977	5,6	9,8	1993	5,9	9,3
Alemania	1984	9,4	8,4	1995	8,5	7,5
Grecia	1974	6,5	5,4	1994	7,5	5,6
Italia	1980	3,9	4,6	1995	6,7	7,1
Irlanda	1987	10,1	10,4	1994	7,8	10,4
Holanda	1979	6,5	9,2	1996	5,3	8,3
Noruega	1983	5,3	6,3	1995	5,5	7,5
Portugal	1982	8,7	12,4	1995	6,7	15,6
España	1990	6,4	8,3	1995	6,7	4,1
Suecia	1981	3,2	6,6	1991	2,4	6,2
Suiza	1992	8,2	10,7	1998	6,3	10,2
Reino Unido	1980	2,5	7,4	1995	4,9	9,7

Fuente: Pereira y Silva-Martins (2002).

dimientos han crecido en la parte superior de la distribución. Un factor que subyace a la distribución de salarios es la habilidad innata dado que individuos menos hábiles deberían predominar en la mitad inferior de la distribución. Así, una explicación para este gráfico es que la educación tiene un impacto mayor sobre los más hábiles que sobre los menos hábiles y esta “complementariedad” entre habilidad y educación parece haber tendido a reforzarse con el transcurso del tiempo<sup>XIV</sup>.

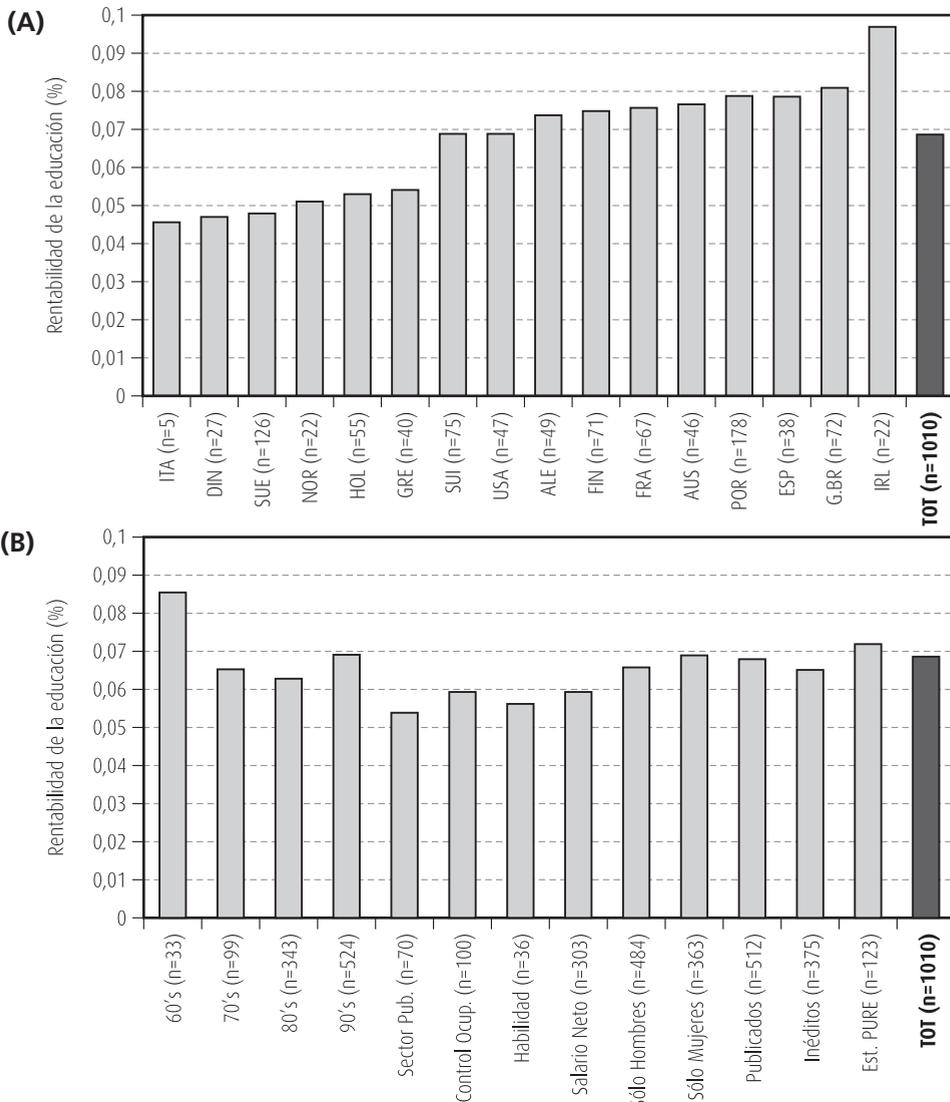
La tabla 3.6, basada en el trabajo del grupo de investigación PURE, posteriormente publicado por Pereira y Silva-Martins (2002), sugiere que en la mayoría de países y para la mayoría de años existe una complementariedad entre educación y habilidad, y que esta es cada vez más fuerte o, al menos, que no se ha debilitado.

### ■ 3.3.7. Metanálisis

Para resumir las diversas cuestiones previamente discutidas seguidamente se va a desarrollar un metanálisis, lo que debe permitir estructurar la revisión de los rendimientos educativos y proporcionar un marco para determinar si nuestras inferencias son sensibles a las especificaciones seleccionadas. Un metanálisis combina e integra los resultados de varios estudios que comparten un aspecto común de modo que sean “combinables” en un sentido estadístico. La metodología es usual en pruebas clínicas en la literatura médica. En la forma

Gráfico 3.5

### METANÁLISIS: ENTRE PAÍSES (A) VARIACIÓN EN LA MUESTRA DE LOS RENDIMIENTOS ESTIMADOS (B)



Fuente: Harmon, Walker y Westergaard-Nielson (2001). Los períodos de tiempo de las clasificaciones 60's/70's/80's y 90's se refieren a los años de datos utilizadas para estimar los modelos. SECTOR PÚBLICO se refiere a estimaciones únicamente para los trabajadores del sector público, CONTROL OCUPACIÓN se refiere a las estimaciones donde se incluye estos controles, HABILIDAD, es donde se controló explícitamente por habilidad en la especificación, SALARIOS NETOS es donde la variable dependiente es el salario neto, en lugar del salario bruto, HOMBRES/MUJERES sólo hacen referencia a especificaciones basadas exclusivamente en muestras de hombres o mujeres, PUBLICADOS/INÉDITOS y ESTIMACIONES PURE, se refieren a estimaciones de rendimientos de la escolarización de estudios actuales publicados y no publicados para diferentes países, y de las estimaciones generadas exclusivamente por el PURE basados en una sencilla especificación escolaridad/edad/edad al cuadrado.

más sencilla el cálculo del rendimiento promedio entre un número de estudios se obtiene ponderando sobre la base del error estándar de la estimación la contribución de cada estudio individual al promedio (véase Ashenfelter, Harmon y Oosterbeek (1999) para más detalles).

En los gráficos 3.5a y 3.5b se presentan los resultados de un metanálisis sencillo basado en la recopilación de tasas de rendimiento de la educación estimadas por MCO del proyecto PURE, complementados con una serie de resultados para los Estados Unidos. Se generaron unas 1010 estimaciones a partir del proyecto PURE<sup>XV</sup> en tres tipos principales de estimaciones de rendimientos educativos —trabajos publicados, trabajos no publicados, y nuevas estimaciones realizadas para el proyecto PURE. Cada bloque se refiere a una muestra diferente de estudios que comparten características (por ejemplo, “EE.UU” indican solo estudios basados en datos de Estados Unidos, “Salarios Netos” indica que la variable dependiente fue el salario neto en lugar del bruto, y “Habilidad” indica que se incluyeron controles por habilidad).

Una serie de cuestiones surgen de este gráfico. A pesar de las cuestiones planteadas anteriormente en este trabajo, existe una considerable similitud en los rendimientos educativos estimados para una serie de posibles datos, con un rendimiento promedio de aproximadamente 6,5% en la mayoría de países y especificaciones del modelo. Hay un conjunto de excepciones a señalar. Se confirma que los países nórdicos generalmente tienen más bajos rendimientos educativos, mientras en el otro extremo los rendimientos para el Reino Unido e Irlanda son superiores al promedio. Además, los rendimientos estimados a partir de estudios para los trabajadores del sector público, y de estudios donde sólo estaba disponible el salario neto (después de impuestos), son en promedio 5%, aproximadamente. Estimaciones en base a muestras de la década de los sesenta también sugieren el mismo resultado.

### ■ 3.3.8. Otras fuentes de variación de los rendimientos: Sobreeducación

Dado el crecimiento en la oferta de trabajadores educados en la mayoría de países de la OCDE en las últimas dos décadas, ha surgido una preocupación en la literatura sobre rendimientos educativos. Si el crecimiento en la oferta de trabajadores cualificados excede al de la demanda, probablemente el resultado sea la sobreeducación en la fuerza laboral. En otras palabras, las cualificaciones que los trabajadores son capaces de aportar exceden a las cualificaciones requeridas. Mason (1996) sugiere que el 45% de los licenciados del Reino Unido desempeñan tareas que no son acordes con las características de los puestos de trabajo que ocupan. La manifestación para el trabajador de esta sobreeducación es un menor rendimiento para los años de educación que exceden los que se necesitan. Con el fin de analizar esta cuestión, el total de años de escolaridad de los individuos debería ser dividido entre años requeridos y exceso de años de educación. La diferencia en los rendimientos para estos dos componentes constituye una manifestación de la sobreeducación.

Existe una serie de vías alternativas para medir la sobreeducación: definiciones subjetivas, basadas en las respuestas de autovaloración a la pregunta directa a los trabajadores sobre si están sobreeducados; o la diferencia entre la escolaridad observada del trabajador

y la escolaridad requerida según el propio encuestado. Claramente se trata de aproximaciones subjetivas que pueden estar sujetas a errores de medida. Por otra parte, la educación requerida para los nuevos trabajadores puede exceder a la de los más antiguos, dado que la falta de experiencia debe ser compensada con una mayor educación. Alternativamente, una medida más objetiva puede ser obtenida al comparar años de escolaridad de los trabajadores con el promedio para la categoría ocupacional en su conjunto. Esta aproximación a menudo es criticada dado que depende de la clasificación de ocupaciones elegida. Según cual sea el grado de detalle de esta clasificación, se pueden mezclar trabajadores en trabajos que requieren diferentes niveles de educación. Además, los niveles de educación requeridos típicamente reflejan el mínimo requerido y no necesariamente reflejan el nivel educativo del candidato deseable.

Groot y Maassen van den Brink (2000) muestran los resultados, a menudo contradictorios, de un metanálisis de los rendimientos educativos y de la sobreeducación (unos 50 estudios en total). Un 26% de los estudios presentan evidencia de que existe una diferencia estadísticamente significativa en los rendimientos para los años de educación requeridos y para el exceso de años. El metanálisis de las regresiones encontró que cuando la sobreeducación se define a partir de la comparación de los años de escolaridad promedio entre categorías ocupacionales, la incidencia de la sobreeducación cae. El rendimiento promedio para los años requeridos de educación es de 7,9%, y este valor aumenta cuando se utilizan los datos más recientes o cuando la educación requerida se define por métodos de autovaloración. El rendimiento promedio de la sobreeducación o, exceso de años frente a los requeridos, es de 2,6%.

Dolton y Vignoles (2000) contrastaron tres hipótesis respecto a la sobreeducación en el mercado laboral de licenciados del Reino Unido en base a la Encuesta Nacional de Licenciados y Diplomados de 1980, que pregunta a los encuestados cual es el requerimiento educativo mínimo para la posición que ocupa actualmente. La primera hipótesis, respecto a que el rendimiento de los años de educación excedente es el mismo que el rendimiento de los años de educación requeridos, resultó claramente rechazada por los datos. Los nuevos licenciados que estaban sobreeducados ganaban considerablemente menos que aquellos en trabajos para licenciados, con una penalización mayor para los trabajos con menor cualificación requerida. La penalización era además más alta para mujeres. La segunda hipótesis es que el rendimiento de la educación excedente, difiere por niveles de titulación. Esta hipótesis se rechaza —aquellos individuos que están sobreeducados con títulos de primer o segundo nivel ganan lo mismo que aquellos sobreeducados con títulos de menor nivel. La última hipótesis era que los rendimientos educativos de los años de sobreeducación diferían entre los sectores público y privado y de nuevo esta hipótesis fue rechazada. Dolton y Vignoles (2000) concluyen que los rendimientos educativos de los años de sobreeducación son menores que los correspondientes a la educación requerida y que la diferenciación sectorial o en el nivel de la titulación no explican este resultado.

Chevalier (2000) analiza la definición de sobreeducación resaltando que licenciados con similar cualificación no son homogéneos en su dotación de competencias, lo que puede

conducir a una sobrestimación del exceso y del efecto de la sobreeducación en los salarios. Utiliza una muestra de dos cohortes de licenciados en el Reino Unido recogida a través de una encuesta postal realizada por la Universidad de Birmingham en 1996, entre graduados de 30 instituciones de educación superior cubriendo el espectro de instituciones del Reino Unido. Fueron seleccionadas cohortes de licenciados de 1985 a 1990, dejando una muestra de 18.000 individuos. Mediante el uso de medidas de satisfacción laboral este estudio separó los considerados “sobreeducados”, en “aparentemente” y “realmente” sobreeducados. El grupo aparentemente sobreeducado percibe una retribución del orden de un 6% menos que los licenciados adecuadamente empleados, pero esta penalización en el pago desaparece cuando se introducen medidas de habilidad. Los licenciados realmente sobreeducados ven reducida la probabilidad de obtener formación en el puesto de trabajo y una penalización en su retribución que puede llegar hasta el 33%. En consecuencia, la sobreeducación genuina parece estar asociada a una falta de competencias que puede explicar entre un 30% y un 40% de la retribución diferencial. Por tanto, lo que usualmente es considerado como sobreeducación representa un fenómeno más aparente que real.

### ■ 3.4. SEÑALIZACIÓN

En la literatura ha dominado la teoría del capital humano y el análisis econométrico ha sido interpretado en este contexto. Un planteamiento alternativo considera que al calcular el valor presente de los costes y beneficios de la educación netos de impuestos y de subsidios, se está aproximando la rentabilidad social a partir de ciertas hipótesis acerca de lo que las tasas brutas de rentabilidad privada pueden representar. (Véase, por ejemplo, OCDE (2001) para consultar los resultados de estos cálculos para distintos países) En realidad, esta rentabilidad social equivale a la privada ajustada por los costes de matriculación e impuestos.

No obstante, una inquietud que se plantea es que la educación puede tener un valor en el mercado laboral no a consecuencia de sus efectos sobre la productividad sino como consecuencia de razones espurias. En particular, la educación puede actuar como una señal de habilidad innata (o de otras características que los empleadores valoran porque ofrecen información relevante sobre la capacidad productiva del empleado). Cabe suponer que los empleadores consideran que la educación alcanzada está correlacionada con la productividad, circunstancia que se dará si los individuos más productivos eligen niveles altos de escolaridad. Un resultado de este tipo operará si el coste en términos de esfuerzo de educarse es más reducido para los individuos más productivos frente a los menos productivos. En estas circunstancias el mercado se caracterizará por alcanzar un equilibrio en el que los individuos más productivos eligen altos niveles de educación. Esta teoría es en gran medida atribuible a Spence (1973), mientras que Riley (2001) ofrece una revisión de la ulterior literatura<sup>xvi</sup>.

Existe una dificultad fundamental en desentrañar la medida en la que la educación constituye una señal de la productividad “preexistente” en contraposición a cual es la contri-

bución de la educación al incremento de la productividad: ambos planteamientos sugieren una correlación positiva entre ganancias y educación pero por muy distintas razones. Ha habido diversas aproximaciones para resolver esta cuestión. Una posibilidad ensayada ha sido la estimación de la relación entre ganancias y educación para los autoempleados, para cuyo colectivo cabe esperar que la educación no tenga valor como “señal” dado que los individuos conocen su propia productividad y no tienen necesidad de autoseñalizarse adquiriendo un mayor nivel educativo (véase Brown y Sessions, 1998). De forma menos convincente también se ha sugerido que los empleados en el sector público pueden percibir salarios que difieren de su productividad a consecuencia de la ausencia de competencia. Naturalmente, en el modelo de señalización las diferencias entre salarios por niveles educativos no necesariamente deben ser un reflejo exacto de las diferencias de productividad sino que es suficiente con que la correlación sea lo bastante elevada para generar auto selección. (véase Psacharopoulos, 1983). En consecuencia las diferencias en los rendimientos de la educación para empleados por cuenta ajena frente a autoempleados, o entre empleados del sector privado frente a los del público es una medida de la educación como señal.

En la tabla 3.7 se ofrecen los resultados obtenidos a partir del *British Household Panel Survey* (BHPS). Los resultados obtenidos a través de la aplicación de MCO permiten obtener tasas de rendimiento comparables e implican un componente de señalización bastante reducido. Un problema potencial con la separación entre autoempleados y empleados por cuenta ajena es que el autoempleo no es aleatorio-individuos con especiales características (frecuentemente no observables) eligen ser autoempleados. En consecuencia, la parte inferior de la tabla muestra los efectos de la educación sobre los salarios al utilizar el método de estimación en dos etapas de Heckman que controla por las diferencias inobservables entre empleados por cuenta ajena y autoempleados. El BHPS contiene información sobre el autoempleo y el valor neto de la vivienda (*housing equity*) características ambas que es probable que estén asociadas con el autoempleo pero que no es probable

Tabla 3.7

### SEÑALIZACIÓN: RENDIMIENTOS PARA ASALARIADOS VERSUS AUTÓNOMOS-BHPS

	ASALARIADOS		AUTÓNOMOS		VALOR SEÑALIZADOR
	RENDIMIENTO	N	RENDIMIENTO	N	
BHPS-MCO	0,0641 (0,002)	10.001	0,0514 (0,008)	1.717	0,0131 (0,012)
Hombres Mujeres	0,1027 (0,002)	9.550	0,0763 (0,015)	563	0,0264 (0,019)
BHPS-Heckman	0,0691 (0,003)	10.001	0,0552 (0,022)	1.717	0,0139 (0,025)
Hombres Mujeres	0,1032 (0,002)	9.550	0,0784 (0,066)	563	0,0248 (0,070)

*Nota:* Entre paréntesis los errores estándar robustos. Los modelos incluyen dummies de año, estado civil, y el número de hijos en tres rangos de edad, dummies regionales, y tasa de desempleo regional. Las estimaciones de modelo de selectividad de Heckman emplea el padre es autónomo, la madre es autónoma, y la equidad en el hogar, como instrumentos.

que muestren correlación con los salarios. Se constata que los resultados, en esencia, permanecen invariables.

El segundo enfoque para diferenciar entre habilidad y productividad es introducir directamente medidas de habilidad. El principal problema con los controles directos de habilidad es que resulta necesario que estas variables no estén afectadas por la educación recibida puesto que en tal caso recogerían el efecto de la educación sobre la productividad. Adicionalmente, las medidas de habilidad deberían hacer referencia la capacidad para ganar dinero en lugar de estar ligadas a tests de inteligencia. Es poco probable que cualquier medida de habilidad satisfaga de forma adecuada estas dos condiciones por lo que se ha abordado este extremo con dos bases de datos específicas.

El *National Child Development Survey* (NCDS) es un estudio por cohortes de todos los individuos nacidos en Inglaterra y Gales en una semana concreta de 1958 cuya evolución fue seguida con especial detalle incluyendo información sobre sus ingresos. Los pertenecientes a esta cohorte fueron sometidos a test de habilidad a las edades de 7, 11 y 16 años. El *International Adult Literacy Survey* (IALS) detalla información sobre ingresos y habilidad referidos ambos al mismo momento en que tiene lugar la entrevista. En el IALS se ofrece información sobre el nivel de alfabetización del individuo referidos a tres escalas: prosa, documentación y cuantitativa. Esta información está referida a la edad que el entrevistado tiene cuando responde al cuestionario.

La tabla 3.8 ofrece estimaciones procedentes del NCDS y del IALS estableciendo diversos controles contemplando variables expresivas de habilidad. En NCDS utilizamos los resultados de tests de Matemáticas e Inglés a la edad de 7 años y se muestran las tasas de ren-

Tabla 3.8

### RENDIMIENTOS EDUCATIVOS POR SEXO EN LAS NCDS E IALS: CONTROLES POR HABILIDAD

		SIN CONTROL POR HABILIDAD	CON CONTROL POR HABILIDAD
NCDS-GB	Mujeres	0,107 (0,007)	0,100 (0,008)
Control a los 7 años	Hombres	0,061 (0,006)	0,051 (0,006)
NCDS-GB	Mujeres	0,107 (0,007)	0,081 (0,009)
Control a los 11 años	Hombres	0,061 (0,006)	0,036 (0,007)
NCDS-GB	Mujeres	0,107 (0,007)	0,071 (0,009)
Control a los 16 años	Hombres	0,061 (0,006)	0,026 (0,007)
IALS-GB	Mujeres	0,106 (0,014)	0,077 (0,013)
Control a la edad actual	Hombres	0,089 (0,009)	0,057 (0,009)

*Nota:* Errores estándar entre paréntesis. Las ecuaciones estimadas incluyen un término cuadrático en la edad, y una tendencia temporal mensual. Los controles por habilidad en las estimaciones con la NCDS, son los puntajes de las pruebas de Inglés y Matemáticas en cuarteles; mientras en las estimaciones con la IALS es el residuo de la regresión de las medidas de habilidad a la edad actual, contra la escolaridad y la edad, para descontar estos efectos.

dimiento de la educación para hombres y mujeres de forma separada. También se comparan estos resultados utilizando controles a las edades de 11 y 16 años. Finalmente, utilizamos información sobre habilidad a partir de la información ofrecida por IALS. Como cabía esperar, la utilización de controles de habilidad a edades más avanzadas mezcla el efecto de la educación sobre la habilidad y en apariencia el sesgo de habilidad aparece incrementado. En consecuencia, los resultados obtenidos a partir de los test de habilidad a la edad de 7 años probablemente ofrezcan la estimación más precisa de la medida en que los rendimientos de la educación están recogiendo el efecto derivado de la habilidad innata. Como puede constatar el efecto es reducido y sugiere un escaso valor de la educación como mecanismo de señalización.

### ■ 3.5. ENDOGENEIDAD DE LA EDUCACIÓN

#### ■ 3.5.1. Aislando el efecto de variación exógena en la escolaridad

Si se quiere conocer como el ingreso del individuo es afectado por un año extra de educación, lo ideal es que se compare el ingreso de un individuo con  $N$  años de educación, con el ingreso del mismo individuo a los  $N-7$  años de escolaridad. El problema de los investigadores es que sólo se observa uno de los niveles de ingreso de interés y el otro es inobservado (Rubin, 1974).

El problema es análogo al encontrado en otros campos, como la ciencia médica: cada paciente recibe un tratamiento eficaz o no, pero observar la efectividad del tratamiento es difícil dado que lo que observamos es el resultado. En estudios médicos la solución habitual a este problema es proveer tratamiento a los pacientes sobre la base de asignación aleatoria. En el contexto de la educación esto es raramente factible, pero puede haber casos en el mundo real, o “experimentos naturales”, que puede ser factible considerar como si se tratase de una asignación aleatoria de individuos a diferentes resultados. La esencia de este enfoque es disponer de un instrumento adecuado para la educación, que no esté correlacionado con los ingresos, y con ello aproximar un ensayo aleatorio, como podría hacerse en un experimento para un estudio clínico.

Una forma muy directa de abordar la cuestión del efecto de un año adicional de educación sobre los salarios, es analizar los salarios de las personas que abandonan la escuela a los 16 años cuando la escolaridad obligatoria se alcanza a los 16, en comparación con los salarios de aquellos que abandonaron la escuela a los 15, justo antes alcanzar la edad mínima. Los datos del FRS son lo suficientemente amplios para permitirnos seleccionar grupos de cohortes pertinentes, y en la tabla 3.9 se muestran los salarios correspondientes.

El efecto tratamiento de haber permanecido en la escuela, aporta la magnitud relevante desde una perspectiva orientada a la implementación de políticas —el efecto de la educación adicional, para aquellos que normalmente no habrían elegido un año extra. Si

Tabla 3.9

## SALARIOS Y EDAD DE ESCOLARIDAD OBLIGATORIA (£/HORA)

	ANTES DE 15 AÑOS PRE-ESCOLARIDAD OBLIGATORIA (1)	ANTES DE 16 AÑOS PRE-ESCOLARIDAD OBLIGATORIA (2)	ANTES DE 16 AÑOS POST-ESCOLARIDAD OBLIGATORIA (3)	% DIFERENCIA ENTRE (3) Y (1) (4)	% DIFERENCIA ENTRE (2) Y (1) (5)
Hombres	7,66	9,56	8,90	14,9	24,8
Mujeres	5,25	6,25	5,81	10,7	19,0

*Nota:* Escolaridad Obligatoria, se refiere al “cambio en la edad de escolaridad obligatoria” de 15 a 16, que ocurrió en 1974.

suponemos que todos aquellos que salen a los 16 de la educación post obligatoria, tienen a los 15 que haber estado en la educación pre-obligatoria, entonces obtenemos un límite inferior para el efecto tratamiento: esto es 14,9% para hombres y 10,7% para mujeres. La cifra anterior es muy similar a la obtenida por Harmon y Walker (1995) utilizando métodos multivariantes más complejos. En cambio, el límite superior del efecto tratamiento es el impacto del año adicional de escolaridad que se ha elegido: esto representa una prima más elevada para hombres, del 24,8% y del 19% para mujeres, lo que refleja el hecho de que las personas que optaron por salir a los 16, son diferentes de aquellos que salieron a los 15 años, en términos de sus otras características.

Más formalmente el grupo tratamiento se elige en un experimento natural, no aleatoriamente, pero independiente de cualquier característica que afecte los salarios. Por lo tanto, no se podría agrupar los datos de acuerdo a la habilidad, pero agrupar por cohorte para captar el efecto antes y después, puede ser válido. La variable que define el experimento natural se puede considerar como una forma de “segmentar los datos” de modo que los salarios y la educación de un grupo puede ser comparado con la de los demás: es decir, se puede dividir la diferencia de salario entre grupos por la diferencia en educación, para formar una estimación de los rendimientos educativos. La restricción clave, es que la variable que define la separación muestral no esté correlacionada con los salarios. Pueden existir diferencias en las variables observables entre los grupos —por ejemplo, el grupo tratamiento podría, ser más alto que el grupo de control— y dado que estas diferencias pueden contribuir a las diferencias en salarios y/o educación, se podrían eliminar tomando las diferencias en el tiempo intra-grupos y restando las diferencias inter-grupos. De aquí que la metodología, frecuentemente, se denomina método de diferencias en diferencias.

Si los datos pueden ser agrupados de modo que las diferencias entre los niveles de educación en los dos grupos sea aleatoria, entonces se puede obtener una estimación de los rendimientos educativos conocida como estimación Wald, al dividir las diferencias en salarios entre grupos por la diferencia en el nivel educativo promedio del grupo. Un ejemplo potencial, es agrupar observaciones de acuerdo al comportamiento fumador en la infancia. El argumento para hacerlo es que fumar en la juventud es señal de tener una alta tasa

Tabla 3.10

### ESTIMACIÓN WALD DE LOS RENDIMIENTOS EDUCATIVOS- AGRUPADOS POR FUMADORES

COMPARA GHS 78-96		FUMADORES (A LOS 16 AÑOS)	NO FUMADORES (A LOS 16 AÑOS)	DIFERENCIA	ESTIMACIÓN WALD
Hombres	Log. Salario	2,36	2,51	0,16	0,16/0,97 = 0,164
	Años Educ.	12,11	13,08	0,97	
Mujeres	Log. Salario	2,01	2,18	0,17	0,17/0,90 = 0,188
	Años Educ.	12,52	13,42	0,90	

de preferencia temporal —dado que fumadores jóvenes revelan que están más dispuestos a incurrir en el riesgo de enfermedad a largo plazo, por el disfrute a corto plazo. El *General Household Survey* de Gran Bretaña ofrece información sobre el hábito de fumar en la juventud para los años 1978 a 1996, y la tabla 3.10 muestra que, examinando estas diferencias entre grupos, los rendimientos estimados de la educación se sitúan alrededor de 16% para hombres y 18% para mujeres.

Una forma estrechamente relacionada de controlar las diferencias en características observables, es controlarlas usando métodos multivariantes. Esta es la esencia del enfoque de variables instrumentales. Es decir, la variable que se usa para agrupar, podría ser usada como una variable explicativa en la determinación del nivel educativo. Esto es útil dado que permite el empleo de métodos multivariantes para controlar por otras diferencias observables entre individuos con diferente nivel educativo. Es además útil, en casos donde la variable es continua —el estudio puede explotar el rango completo de variación en los instrumentos, más que simplemente emplearlo para categorizar los individuos en dos (o más) grupos. Por medio de la explotación de instrumentos para la educación, que no estén correlacionados con los ingresos, el enfoque de variables instrumentales generará estimaciones insesgadas de los rendimientos educativos.

Consideremos el modelo  $\log w_i = X_i\beta + \gamma s_i + u_i$  donde  $s_i = Z_i' \alpha + v_i$ . La estimación de la ecuación del logaritmo de salarios por MCO aportará una estimación insesgada de  $\beta$  si, y solo si,  $s_i$  es exógeno, es decir, cuando no existe correlación entre los dos términos de error. Si esta condición no es satisfecha, se deben emplear métodos de estimación alternativos, dado que la estimación MCO será sesgada. La correlación puede ser no nula, porque algunas variables importantes relacionadas con la escolaridad y los ingresos se han omitido del vector  $X$ . Medidas de motivación o habilidad, como el cociente intelectual, son posibles ejemplos. Es importante tener en cuenta, que incluso una lista muy amplia de variables incluidas en el vector  $X$  no será exhaustiva, de modo que las estimaciones de los rendimientos educativos basados en MCO no darán el efecto causal de la escolaridad sobre los salarios<sup>xvii</sup>, en el sentido que el coeficiente  $\beta$  capta algo del efecto que en otro caso sería atribuido a las variables omitidas. Por ejemplo, si la variable omitida es la motivación, y si

la educación y los salarios están positivamente correlacionados con la motivación, la estimación MCO ignora que es más probable que individuos más motivados ganen más que los menos motivados, incluso para similares niveles educativos.

Con el fin de modelar la relación entre escolaridad y salarios, deberíamos utilizar la ecuación de educación para estimar el valor predicho de la escolaridad. Entonces reemplazamos la educación en la función de salarios por su nivel predicho. Como la escolaridad predicha esta correlacionada con la escolaridad observada, esta variable captará el efecto de la educación en los salarios. No obstante, no hay razón para suponer que la escolaridad predicha esté correlacionada con el término de error en la ecuación de ingresos, por lo que el rendimiento estimado basado en la escolaridad predicha será insesgado. Este es el método de estimación por Mínimos Cuadrados en dos Etapas, que es un caso especial del método de variables instrumentales (o VI) y que capta su esencia.

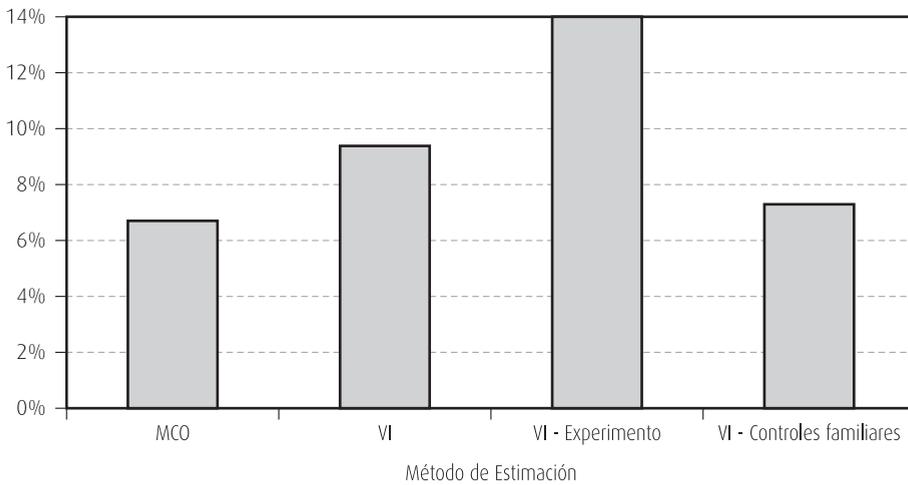
La dificultad en este procedimiento es de “identificación”. Para identificar o aislar el efecto de la escolaridad sobre los ingresos, debemos centrar la atención sobre las variables contenidas en el vector  $Z_i$  que no están incluidas en  $X_i$  (véase la discusión en Heckman (1990) para más detalles). Es decir, debe existir una variable que sea un factor determinante de la escolaridad, que pueda razonablemente ser omitida en la ecuación de salarios. En esencia, esto equivale a examinar de qué manera los salarios difieren entre grupos, cuya educación es diferente por razones exógenas. Por ejemplo, algunos individuos se enfrentan a una escolaridad obligatoria que no coincidió con la que se encontraron otros, o podrían haber empezado la escuela más temprano por otras razones aleatorias (es decir, por razones que no están correlacionadas con los salarios que finalmente obtuvo), como fumar en la juventud, que como antes se sugirió, está asociado con la tasa de preferencia temporal.

### ■ 3.5.2. Resultados de los Estudios con Variables Instrumentales-Evidencia Internacional

En el gráfico 3.6 presentamos los resultados de un metanálisis de estudios que abordan la endogeneidad de la educación basado en los resultados obtenidos con las bases de datos del proyecto PURE, presentados anteriormente. En comparación con un promedio de 6.5% (n=863) para la estimación por MCO, observamos mayores rendimientos educativos en los estudios con VI de forma generalizada (de aproximadamente 9%, n = 79), y en particular en los estudios de VI basados en reformas educativas (entre 13% a 14%, n = 17) —parecido a la estimaciones de diferencias en diferencias presentadas anteriormente. En cambio, los estudios con VI que utilizan el *background* familiar como instrumento tienen en promedio rendimientos similares a la estimación MCO. En los pocos ejemplos donde se ha contrastado la validez del *background* familiar como instrumento, se ha demostrado que es un instrumento débil (Rischall, 1999).

Gráfico 3.6

## METANÁLISIS DE MODELOS DE ENDOGENEIDAD EN LA EDUCACIÓN



Fuente: Harmon, Walker y Westergaard-Nielsen (2001). Número de estimaciones = 942.

La tabla 3.11 describe algunos de los resultados de estudios clave en la literatura. Angrist y Krueger (1991) utilizan la presencia de cambios en la ley de escolaridad obligatoria en Estados Unidos, y el trimestre del año en el que la persona nació, como instrumentos. La idea subyacente es que una persona que ha nacido a principios del año (en el primer trimestre) alcanza la edad mínima de escolaridad obligatoria después, con un menor número de meses de educación, respecto a las personas nacidas a finales del año. La cantidad de educación alcanzada, está directamente relacionada con el trimestre de nacimiento, mientras al mismo tiempo parecen no existir razones para creer que el trimestre de nacimiento tenga un efecto independiente sobre los ingresos. La estimación directa por MCO ofrece una estimación de los rendimientos educativos de 0,063, mientras con el método de VI se obtiene una estimación de 0,081<sup>XVIII</sup>.

En otro estudio, Angrist y Krueger (1992) explotan la idea de que dado que entrar en la universidad representó una reducción en la probabilidad de participar en la guerra de Vietnam, ello constituyó un efecto exógeno sobre la demanda de escolaridad. El instrumento fue el número aleatorio dado a los individuos en el sorteo empleado para reclutar jóvenes para la guerra de Vietnam. De nuevo los resultados por VI fueron más elevados que los obtenidos por MCO pero sin que la diferencia resultase significativa. Como un trabajo posterior reflejó (véase Bound *et al.*, 1995), esta ausencia de significatividad de la diferencia puede ser una consecuencia de que el instrumento utilizado apenas resultaba marginalmente significativo en sus efectos sobre la decisión de educarse.

Tabla 3.11

## ESTUDIOS SOBRE VARIABLES INSTRUMENTALES

ESTUDIO	MUESTRA	MCO	VI	INSTRUMENTOS
Angrist y Krueger (1991)	EE,UU, 1970/1980 Censo: Hombres nacidos 1920-29, 1930-39, 1940-49	0,063 (0,000)	0,081 (0,033)	Año · Trimestre de nacimiento; Estado · Trimestre de nacimiento
Angrist y Krueger (1992)	EE,UU, 1979-85 CPS: Hombres nacidos 1944-53 (reclutas potenciales para La guerra de Vietnam)	0,059 (0,001)	0,066 (0,015)	Número lotería servicio militar · Año de nacimiento
Card (1995)	EE,UU, NLS: Hombres entre 14-24 años, en 1966, empleados en la muestra de 1976,	0,073 (0,004)	0,132 (0,049)	Cercanía a la universidad País de residencia en 1966,
Butcher y Case (1994)	EE,UU, PSID 1985: Mujeres blancas con 24 años o más	0,091 (0,007)	0,185 (0,113)	Presencia de hermanos (hermanas)
Uusitalo (1999)	Fuerzas básicas de Defensa de Finlandia Datos de prueba de habilidad con Registros de impuesto a la renta en Finlandia	0,089 (0,006)	0,129 (0,018)	Ingreso de los padres y Educación, lugar de residencia
Meghir y Palme (1999)	Suecia-Hombres	0,028 (0,007)	0,036 (0,021)	Reformas al currículo sueco
Duflo (1999)	Indonesia-Hombres	0,077 (0,001)	0,091 (0,023)	Proyecto construyendo escuelas de Indonesia
Denny y Harmon (2000)	Irlanda-datos del ESRI 1987- <i>Hombres</i>	0,080 (0,006)	0,136 (0,025)	Reforma escolar en Irlanda Eliminación de las tasas para la escolaridad secundaria

Tabla 3.11 (continuación)

## ESTUDIOS SOBRE VARIABLES INSTRUMENTALES

ESTUDIO	MUESTRA	MCO	VI	INSTRUMENTOS
Dearden (1998)	Reino Unido NCDs: <i>Hombres</i>	0,048 (0,004)	0,055 (0,005)	Composición familiar, Educación de los padres
Harmon y Walker (1995)	Reino Unido FES 78-86, <i>Hombres</i>	0,061 (0,001)	0,152 (0,015)	Cambios en la escolaridad Obligatoria
Harmon y Walker (1999)	Reino Unido GHS 92, <i>Hombres 16-64</i>	0,049 (0,000)	0,140 (0,005)	Cambios en la escolaridad Obligatoria y Reformas educativas
Harmon y Walker (2000)	Reino Unido NCDs: <i>Hombres</i>	0,050 (0,005)	0,099 (0,019)	<i>Background</i> familiar
Pons y Gonzalo (2001)	España: <i>Hombres 16-64</i>	0,064 (0,004)	0,107 (0,010)	Intervención a la Política Educativa <i>Background</i> familiar Trimestre de nacimiento

*Nota:* Errores estándar en paréntesis. Véase Card (2001) para información adicional.

Card (1995) emplea la distancia a la universidad como instrumento de la escolaridad, basado en la observación de un mayor nivel educativo para hombres que crecieron cerca de una universidad. Obtiene rendimientos del 13,2%, comparados con las estimaciones MCO cercanas al 7%. Sin embargo, de nuevo las estimaciones eran bastante imprecisas. Butcher y Case (1994), en uno de los pocos ejemplos basados en una muestra de mujeres, encontraron que las mujeres que crecieron en hogares con una hermana obtuvieron menos educación que quienes crecieron sólo con hermanos. Empleando esto como instrumento, de nuevo encontraron que las estimaciones por VI superaban a las estimaciones por MCO, y de hecho los rendimientos estimados fueron más del doble en este estudio. Uusitalo (1999) emplea el hecho de que todos los hombres finlandeses deben prestar servicio militar donde realizan un test de aptitud. Al cruzar estos datos con los registros del impuesto sobre la renta, el estudio estima ecuaciones de ingresos para hombres, basadas en instrumentos de *background* familiar. Los resultados de nuevo sugieren una estimación mayor para VI frente a MCO, del orden de un 45%, además, estadísticamente significativa.

Un enfoque un tanto diferente es empleado en el documento de Duflo (1999) donde la estimación se basa en la participación de individuos en un programa masivo de inversión en educación en Indonesia a principios de los años setenta. Los individuos fueron asignados al programa en función de su cohorte de nacimiento, y la intensidad del tratamiento dependió del distrito donde vivía (las inversiones fueron dirigidas a regiones donde las matriculas eran históricamente bajas). Meghir y Palme (1999) siguen una estrategia similar en su análisis de las reformas en Suecia en los años cincuenta, que pretendían aumentar la escolaridad a nivel nacional. Se realizó una prueba piloto en varios distritos escolares antes de su implementación a escala nacional, y fue de este experimento previo donde se obtuvo la variación en la escolaridad. Ambos estudios se basan en reformas a gran escala, que pueden ser vistas como “experimentos naturales” ya que su efecto difiere entre los individuos. Igualmente, Denny y Harmon (2000) utilizan el cambio estructural en el sistema educativo Irlandés en los años sesenta, que afectó el conjunto de la población en edad escolar en un sentido que difirió según *background* socioeconómico. Finalmente, Pons y Gonzalo (2001) estiman los rendimientos de la educación para España usando instrumentos basados en las intervenciones en la política educativa, variables de *background* familiar y trimestre de nacimiento como instrumentos, y hallan que los rendimientos educativos aumentaron de 6.4% estimados por MCO, a 10.7% estimados por VI<sup>XIX</sup>.

Hay un pequeño número de ejemplos en la literatura del Reino Unido empleando este enfoque, que también se resumen en la tabla 3.11. Dearden (1995, 1998) repite la idea de Butcher y Case (1994), utilizando la presencia de hermanos como instrumento de la escolaridad. Esta investigación empleó los datos del *National Child Development Study* (NCDS), y encontraron aumentos en los rendimientos educativos estimados comparados con los correspondientes MCO. En una serie de estudios Harmon y Walker (1995, 1999, 2000), utilizan como instrumentos los cambios en las leyes de escolaridad obligatoria en los años cincuenta y setenta, entre otras reformas educativas (como la creación de ayudas a la manu-

tención para estudiantes). Para diferentes bases de datos del Reino Unido, aparece como resultado robusto que las estimaciones de los rendimientos por VI fueron significativamente mayores, que las MCO que se sitúan en el rango del 5-7% por año de educación.

Para UK las diferencias entre VI y MCO son claramente grandes y resultan acordes con la experiencia internacional. Aunque estos resultados por VI coinciden con las sencillas estimaciones de Wald, presentadas al principio, no obstante, es importante someter el resultado a un examen más detallado. En la tabla 3.12 presentamos resultados para un conjunto de bases de datos y especificaciones que utilizan el hábito de fumar como instrumento, en el sentido propuesto por Evans y Montgomery (1994), donde se argumenta que fumar sugiere una preferencia temporal especialmente alta: es decir, las altas tasas de descuento para individuos fumadores indican que conceden un peso importante a satisfacer sus deseos actuales a costa del futuro. Se supone que fumar a los 16 años no está correlacionado con los ingresos actuales, pero está correlacionado con la elección educativa. En la tabla podemos ver mayores rendimientos estimados para VI, frente a los resultados por MCO para *General Household Survey* del Reino Unido (GHS). Cuando se usó fumar actualmente, se

Tabla 3.12

### RESULTADOS ADICIONALES - EL HÁBITO DE FUMAR COMO INSTRUMENTO

DATOS E INSTRUMENTOS	HOMBRES		MUJERES	
	RENDIMIENTOS ESTIMADOS	N	RENDIMIENTOS ESTIMADOS	N
GHS: MCO	0,064 (0,002)	14.424	0,092 (0,002)	11.759
GHS: VI -Fuma Actualmente	0,205 (0,012)	14.424	0,163 (0,011)	11.759
GHS: VI -Fuma a los 16 años	0,095 (0,007)	17.907	0,126 (0,008)	17.047
BHPS: MCO	0,064 (0,002)	8.284	0,103 (0,002)	8.987
BHPS: VI - Fuma Actualmente	0,209 (0,014)	8.284	0,168 (0,011)	8.987
NCDS: MCO (sin controles familiares)	0,075 (0,005)	3.169	0,120 (0,006)	2.319
NCDS: VI - Fuma Actualmente (sin controles familiares)	0,203 (0,029)	3.161	0,241 (0,030)	2.736
NCDS: VI - Fuma a los 16 años (sin controles familiares)	0,084 (0,030)	2.486	0,219 (0,025)	2.150
NCDS: MCO (con controles familiares)	0,061 (0,006)	3.169	0,107 (0,007)	1.981
NCDS: VI - Fuma Actualmente (con controles familiares)	0,191 (0,031)	2.311	0,215 (0,043)	1.978
NCDS: VI - Fuma a los 16 años (con controles familiares)	0,080 (0,033)	1.972	0,207 (0,032)	1.692

*Nota:* Entre paréntesis los errores estándar robustos. Los modelos incluyen *dummies* de año, estado civil, y el número de hijos en tres rangos de edad, *dummies* regionales, y tasa de desempleo regional. El número de observaciones difiere por la existencia de *missing* en algunas variables.

obtuvieron rendimientos muy altos, frente a incrementos más moderados cuando se empleó fumar a los 16 años. Esto es probablemente porque fumar actualmente está correlacionado con el nivel de ingreso actual, lo que invalida fumar actualmente como variable instrumento. Una crítica de emplear fumar a los 16 años como VI, es que puede estar correlacionado con los salarios actuales vía su asociación con el *background* de los padres. Así, en el último bloque de la tabla 3.12 además se controla por *background* familiar y se siguen obteniendo unos rendimientos similares a los del bloque anterior sin controles familiares.

### ■ 3.5.3. ¿Por qué son las estimaciones por VI mayores a las MCO?

En el modelo de Card (1999), se permite que los rendimientos educativos varíen entre individuos, y el rendimiento marginal de la educación es función decreciente de la educación. Cuando el instrumento se construye sobre la base de la pertenencia al grupo tratamiento, la estimación por VI de los rendimientos educativos es la diferencia en la esperanza del logaritmo de los ingresos entre el grupo control y el grupo tratamiento, dividido por la diferencia en la escolaridad esperada para los dos grupos. Esto implica que si todos los individuos en la población tienen el mismo rendimiento marginal, la estimación por VI, es una estimación consistente de la tasa de rendimiento marginal promedio. Sin embargo, si se permite que el rendimiento educativo cambie entre individuos, la estimación por VI es un rendimiento ponderado, donde los pesos reflejan la medida en la cual un subgrupo es afectado por el tratamiento o instrumento. Si solo un subgrupo es afectado por la intervención, el estimador por VI, ofrecerá la tasa de rendimiento marginal para *ese subgrupo*. Un estudio similar de Lang (1993), considera además este problema en el contexto de heterogeneidad de las diferencias en las tasas de descuento.

Dada esta interpretación, queda claro que el estimador por VI puede exceder al obtenido por MCO si la intervención afecta a un subgrupo con rendimientos marginales de la educación relativamente elevados. En el modelo de Card (1999), esto es posible en la medida que pequeños niveles de educación pueden implicar altos rendimientos de la escolaridad si la variabilidad de la habilidad es reducida en comparación a la variabilidad de la tasa de descuento individual. Si la intervención afecta a los que tienen niveles de escolaridad inferiores a la media, la estimación por VI ofrecerá valores más elevados que los resultados “promedio” por MCO, reflejando la mayor tasa de descuento de quienes tienen menor escolaridad. Esto se ha sugerido como justificación de los resultados de, por ejemplo, Angrist y Krueger (1991) con relación a las leyes de la escolaridad obligatoria, y es un ejemplo concreto del problema más general de estimar efectos para grupos marginales afectados por el tratamiento conocido como Efecto Tratamiento Local Promedio (o ETLP —véase Imbens y Angrist, 1994).

Además, si el instrumento(s) está correlacionado con el verdadero valor de la educación, pero no con el error de medida de la educación, el enfoque VI puede ser empleado, y

la presencia de errores de medida no debería afectar la estimación de los rendimientos educativos por VI. Lo que variará es la interpretación que se puede dar a la diferencia entre los resultados MCO y VI. Como tal, la diferencia puede ser ahora atribuida al efecto combinado del error de medida y de la endogeneidad de la educación. El estudio de Ashenfelter y Krueger (1994), estima la ratio de fiabilidad de los años de escolaridad en el 90% (la ratio de varianza del error de medida respecto a la varianza total de la educación), sugiriendo que aproximadamente el 10% del total de la varianza en la escolaridad es debido al error de medida. Además, Uusitalo (1999) emplea información sobre escolaridad utilizando datos de registro que se actualizan empleando información de los propios centros de enseñanza, por lo que el grado de error de medida debe ser con toda probabilidad mucho menor. A pesar de que el error de medida debe ser un problema menor en estos estudios, ambos encuentran grandes y significativos sesgos a la baja en los estimadores por mínimos cuadrados. Atendiendo a esta evidencia, los errores de medida aparecen como un candidato poco probable para explicar la diferencia VI/MCO.

#### ■ 3.5.4. Relevancia del instrumento y validez del instrumento

Bound *et al.* (1995), recomiendan cautela en el uso de VI. Las VI pueden ser consideradas como una vía para dividir la varianza de la escolaridad entre un componente endógeno y otro exógeno. Ello se consigue al incluir una variable (o variables) en la ecuación para explicar las decisiones de educarse, que no esta (están) en la ecuación de salario. La consistencia del estimador por VI depende del supuesto de que las variables instrumentales estén correlacionadas con las decisiones de escolarizarse de los individuos, pero que no lo estén con los salarios obtenidos. Si existe una relación entre instrumentos y salarios, la estimación por VI puede conducir a importantes inconsistencias. Por otra parte, una relación débil entre escolaridad e instrumentos se añadirá a este problema. Como ejemplo, Bound *et al.* (1995) reestiman las ecuaciones de Angrist y Krueger (1991), encontrando que muchos de los instrumentos utilizados en su estudio estaban prácticamente no correlacionados con la escolaridad lo que puede ocasionar que el sesgo derivado del empleo de VI exceda al que se desprende de la aplicación de MCO.

En experimentos naturales, las asignaciones no aleatorias a los grupos tratamiento y control, da lugar a similares problemas de inconsistencia. Card (1999) señala que el estudio de Harmon y Walker (1995) divide la muestra de acuerdo con uno de los tres cambios posibles de escolaridad obligatoria: las personas nacidas entre 1933 y 1958 se consideraron como grupo de control y aquellos nacidos antes de 1933 o después de 1958 fueron los grupos tratamiento sobre la base de la introducción de dos cambios en la enseñanza obligatoria. No obstante, las cohortes mayores pueden ser diferentes en otros sentidos —en particular el nivel educativo de la cohorte nacida antes de 1933 puede estar afectada por la Segunda Guerra Mundial (véase también, por ejemplo, Ichino y Winter-Ebmer, 2000).

Finalmente, Ashenfelter, Oosterbeek y Harmon (1999) señalan la existencia del denominado sesgo de publicación. El rendimiento educativo promedio en un metanálisis de la rentabilidad de la educación estimada por MCO es del 6%, frente a un promedio de más de 9% de la estimación por VI. Ashenfelter *et al.* modelizan la probabilidad de ser observado en una muestra de rendimientos publicados como función creciente del nivel de significatividad de la diferencia entre las estimaciones por MCO y VI. En otras palabras, resultados más significativamente distintos entre ambas estimaciones tienen una mayor probabilidad de aparezcan en una muestra de rendimientos publicados. Cuando la contribución de este factor se cuantifica se constata que aproximadamente explica del orden de dos terceras partes de la diferencia entre MCO y VI.

### ■ 3.5.5. Evidencia adicional-estimadores de efectos fijos y métodos de correspondencia

#### ■ 3.5.5.1. *Estimadores panel*

Las técnicas de datos de panel se pueden emplear para abordar la heterogeneidad tratando la heterogeneidad inobservada como un efecto fijo. Se supone que los efectos inobservables son constantes en el tiempo, y por lo tanto, las observaciones del mismo individuo en diferentes periodos de tiempo, aportarán la información necesaria para eliminar los efectos inobservables. No obstante, la aplicabilidad de las técnicas de panel para evaluar los rendimientos educativos es limitada, dado que sólo tenemos información de ingresos una vez finalizado el período de escolarización. Tomar primeras diferencias de los ingresos elimina no solo el efecto fijo inobservado, sino también la información sobre educación. Se requiere, por tanto, información sobre ingreso de los individuos antes y después de educarse, y como tal, esta información sólo estará disponible para aquellos individuos que posteriormente retoman los estudios. Angrist y Newey (1991) encuentran que cerca del 19% de los varones ocupados encuestados en la *National Longitudinal Survey of Youths* (NLSY, un estudio de cohorte realizado en Estados Unidos que sigue personas jóvenes a través del tiempo), declaran un mayor nivel de escolaridad en las últimas olas de datos, en cuyo caso la educación no podría asimilarse a un efecto fijo<sup>xx</sup>. No obstante, no existen ejemplos adicionales en esta literatura.

#### ■ 3.5.5.2. *Datos de gemelos o hermanos*

La tabla 3.13 presenta algunos hallazgos recientes de la literatura basada en muestras de hermanos o gemelos. Este enfoque explora la hipótesis de que los hermanos son más parecidos que una selección aleatoria de un par de individuos, dado que comparten la misma herencia, soporte financiero, influencia de los compañeros, influencias geográficas y sociológicas, etc. Esta literatura intenta eliminar el sesgo de habilidad estimando la ecuación

Tabla 3.13

### ESTUDIOS SOBRE RENDIMIENTOS EDUCATIVOS CON INFORMACIÓN DE GEMELOS/HERMANOS

AUTOR	DATOS	MCO	VI
Ashenfelter y Rouse (1998)	Encuesta gemelos de Princeton	7,8%	10%
Rouse (1999)	Encuesta gemelos de Princeton	7,5%	11%
Miller <i>et al.</i> (1995)	Registro Australiano de gemelos	4,5%	7,4%
Isacsson (1999)	Gemelos Suecos del mismo sexo	4%	5,4%
Ashenfelter y Zimmermann (1997)	Hombres jóvenes de la NLS	4,9%	10%
Bonjour, Haskel y Hawkes (2000)	Unidad de Investigación de niñas Gemelas de Santo Tomas	6,2%	7,2%

ción de salarios en forma de diferencias entre hermanos. El método se fundamenta en suponer que al calcular diferencias se elimina el efecto de la habilidad innata. En efecto, si la variable omitida es la habilidad y esta habilidad innata es la misma para los dos hermanos, al estimar la ecuación en diferencias se eliminará el sesgo asociado al carácter no observable de la variable.

La revisión de estudios sobre gemelos de Griliches (1979) concluye que el rendimiento educativo estimado, cuando el sesgo de habilidad se elimina a través de la diferenciación entre pares de gemelos, es menor que el rendimiento estimado para la muestra completa (es decir, sin diferenciar). Bound y Solon (1998), señalan que los datos de gemelos en Estados Unidos parecen tener grandes diferencias en escolaridad, superiores a las que se darían en el emparejamiento aleatorio de individuos no relacionados, lo que permite albergar dudas sobre los datos. No obstante, una crítica más profunda a este enfoque se ha centrado en los supuestos subyacentes. Si la habilidad tiene un componente individual, además de un componente familiar que no es independiente de la variable escolaridad, el enfoque intra familiar no podría aportar estimaciones menos sesgadas. El problema del error de medida puede ser particularmente nocivo en métodos basados en diferenciación, por ejemplo en la literatura de gemelos, dado que al tomar diferencias, los efectos sobre el sesgo derivados del error de medida se acrecientan.

Contribuciones recientes a la literatura de gemelos han tratado de abordar el problema del error de medida instrumentando la educación de, por ejemplo, el gemelo *A*, a través de la educación declarada por el gemelo *B*. Ashenfelter y Krueger (1994) recogieron datos en el festival anual de gemelos en 1991, y encontraron evidencia contra el supuesto convencional de que la estimación MCO presenta sesgo al alza y a favor de los resultados encontrados en la literatura de VI. Sin embargo, al introducir correcciones por error de medida en los niveles de escolaridad reportados por el individuo, se obtienen estimaciones de los rendimientos mucho más elevadas, del orden de 12-16%. La posible no aleatoriedad de

las bases de datos y las relativamente pequeñas muestras empleadas, han llevado a críticas, pero los resultados de Ashenfelter y Zimmerman (1997), basados en datos de hermanos apoyan este resultado. Más aún, Miller *et al.* (1994), emplean una muestra mucho más grande de gemelos de una encuesta australiana y, aplicando la misma técnica de Ashenfelter y Krueger (1994) encuentran evidencia fuerte de sesgo a la baja en las estimaciones por mínimos cuadrados. El único estudio para el Reino Unido ha sido efectuado por Bonjour *et al.* (2000), y aunque se trata de una muestra de mujeres que participan en un panel de salud, los autores comprueban que la muestra coincide bastante bien con la *Labour Force Survey* (LFS). Además, la instrumentación de las diferencias entre gemelos no parece sugerir problemas de endogeneidad.

La principal debilidad de todos estos estudios es la escasa o nula atención que se le ha dado al por qué los gemelos tienen diferentes niveles de educación, con la excepción del estudio de Bonjour *et al.* (2000). La literatura supone que las diferencias en educación se producen de forma aleatoria, y no es evidente que éste sea el caso. Si no es el caso, entonces la literatura de gemelos presenta precisamente el mismo problema de endogeneidad que ha afectado al resto de la literatura.

#### ■ 3.5.5.3. *Métodos de correspondencia*

En su revisión de la literatura Blundell *et al.* (2001) consideran la estimación de los rendimientos educativos usando variables instrumentales, enfoques de función control y los denominados *propensity score matching methods*. Utilizando el panel del NCDS, encontraron que las estimaciones de los rendimientos por MCO eran significativamente menores a las obtenidas por VI o por el enfoque de funciones control. En base a los métodos de correspondencia, los rendimientos estimados se situaron entre las estimaciones por MCO y VI al considerar el efecto tratamiento para los tratados, y por debajo de la estimación por MCO y VI, cuando se considera el efecto tratamiento para los no tratados.

### ■ 3.6. CONCLUSIONES

A pesar de un fundamento teórico bien desarrollado, la estimación de los rendimientos de la educación ha sido objeto de un debate considerable en la literatura económica. Una característica dominante en la literatura que estima la función de ingresos de capital humano es que la educación es endógena, y ello ha centrado los esfuerzos de investigación recientes. Con respecto a los rendimientos educativos se puede extraer un conjunto de conclusiones.

El análisis simple de ingresos medios para diferentes niveles de educación puede enmascarar una serie de cuestiones. La omisión de controles adicionales supone que las variables inobservadas que afectan a los salarios están no correlacionadas con la educación

—lo que parece poco realista. El análisis de regresión multivariante basado en microdatos del Reino Unido sugiere un rendimiento educativo entre un 7% y 9% para hombres, y entre un 9% y 11% para mujeres, cuando se controla por educación y experiencia. Este parece ser el límite superior de los rendimientos educativos para Europa, donde los países nórdicos en particular tienen una baja rentabilidad promedio de la educación. Los rendimientos educativos son relativamente insensibles a cambios en la especificación MCO (como incluir controles por estado civil, tamaño del hogar o estar sindicado), pero algunas diferencias son dignas de mención. La utilización de distintas medidas de experiencia (basadas en la experiencia declarada o en la experiencia potencial calculada sobre la base de la diferencia entre la edad y la edad a la que se dejó la escuela) tenderá a aumentar el rendimiento educativo en aproximadamente un 1%. Incluir controles por posición ocupacional tenderá a tener el efecto opuesto, reduciendo los rendimientos alrededor de un 1%. Basar la estimación en muestras de personas ocupadas puede, además, sesgar al alza la estimación de los rendimientos, al menos para muestras de mujeres, pero la evidencia sugiere que este efecto es pequeño, a pesar de ser significativo.

La especificación básica supone que el logaritmo del salario es lineal respecto a la educación, de modo que cada año de escolaridad añade la misma cantidad porcentual de salario con independencia de la naturaleza del año de educación considerado. Existe cierta evidencia de que algunos años de educación comportan una prima o una penalización diferente. Por ejemplo, abandonar la escuela el año después de obtener un título, podría generar una menor tasa de rendimiento para ese año, generando un cambio de pendiente en el perfil ingreso-educación.

Dado el incremento en la oferta de trabajadores educados en la mayoría de países de la OCDE existe una preocupación en la literatura, con respecto a que la capacitación de los trabajadores sea superior a los requisitos que el puesto de trabajo exige. Esto se manifiesta en un menor rendimiento educativo para los años de educación que exceden a los requeridos. Uno de los principales problemas de que adolece esta literatura es la deficiente definición de sobreeducación en las bases de datos disponibles, generalmente basada en la propia respuesta del individuo entrevistado. En los trabajos basados en la satisfacción laboral se ofrece una definición más completa, y el aparente efecto negativo de la sobreeducación se elimina cuando se controla por habilidad, pero cuando la sobreeducación es auténtica, la penalización parece ser mayor de lo esperado. Esta es una importante consecuencia del cambio en la calidad de los graduados que genera el sistema educativo. En primer lugar, el título no es suficiente para asegurar un trabajo como licenciado —se valoran otras habilidades complementarias de los licenciados. En segundo lugar, ya que puede surgir auténtica sobreeducación, es evidente que el mercado laboral no se ajusta lo suficientemente rápido. Se puede requerir un cierto grado de planificación de los recursos humanos para asegurar que no se dé una sobreoferta de tipos específicos de licenciados.

Los rendimientos educativos pueden diferir en la distribución salarial. Evidencia basada en los métodos de regresión cuantílica sugieren que los rendimientos son mayores para

aquellos en el decil superior de la distribución de ingresos comparado con aquellos en el decil inferior. Por otra parte, esta desigualdad puede haber aumentado en años recientes. Una explicación para este fenómeno es la complementariedad entre habilidad y educación —si una persona más hábil gana más, esto podría explicar los mayores rendimientos en los deciles superiores de la distribución salarial. Este hallazgo tiene importantes consecuencias para la educación, y la política impositiva y de seguridad social: un ejemplo (posiblemente extremo) es que los bajos rendimientos de la inversión en individuos poco hábiles, y los altos rendimientos de la inversión en individuos más hábiles, implica que la inversión educativa debería estar sesgada hacia los individuos más hábiles. La desigualdad resultante podría posteriormente abordarse a través de impuestos redistributivos y políticas de seguridad social.

Es posible que los rendimientos educativos de hecho reflejen la habilidad subyacente que la educación señala —en otras palabras, la educación es una señal de la productividad inherente del individuo, más que un medio para aumentar la productividad. Las estimaciones presentadas aquí del componente señalizador de los rendimientos sugieren que el efecto es bastante pequeño. De acuerdo con los resultados obtenidos a partir de las bases de datos donde se dispone de mediciones directas de la habilidad, la inclusión de tales medidas reduce los rendimientos educativos al menos en un punto porcentual. Esta reducción tiende a ser mayor cuando las medidas de habilidad se toman a una edad más avanzada, lo que puede ser un reflejo de que cuanto mayor es la edad del individuo, mayor es la probabilidad de que la habilidad medida refleje en parte los efectos de la escolaridad. Si deseamos medir los rendimientos educativos, lo ideal sería comparar los ingresos de un mismo individuo con dos niveles distintos de escolaridad, pero en la práctica solo se observa un nivel de escolaridad por individuo. La literatura reciente intenta abordar este problema encontrando “experimentos” en la economía, que aleatoriamente asignen grupos de individuos a diferentes niveles de escolaridad, permitiendo usar VI. El efecto del cambio en el procedimiento de estimación puede ser considerable. En general, los rendimientos promedio de la escolaridad a partir de MCO se sitúan alrededor de un 6% pero ascienden al entorno del 9% si se emplean VI. El Reino Unido ocupa el extremo superior del rango internacional, con rendimientos que oscilan entre el 7% y 9% en caso de MCO y el 11% ó 15% si se emplean VI. Una limitación de esta metodología es que los mayores rendimientos pueden reflejar el rendimiento para el subgrupo particular afectado por la política de intervención. Así, por ejemplo, cambios en la educación obligatoria pueden afectar a aquellos individuos que conceden un menor *valor* a la educación, por lo que las estimaciones del rendimiento educativo, basadas en esos cambios pueden ser la estimación de los rendimientos para ese grupo. En resumen, se debe tener cuidado en la interpretación de los rendimientos educativos estimados por VI, como un indicador de los rendimientos para todos los individuos, sin un conocimiento detallado de los efectos de las intervenciones utilizadas en la estimación de los rendimientos.

Una limitación adicional es que generalmente la intervención suele tener un efecto débil sobre la escolaridad, y que esta falta de contenido informativo del instrumento puede

introducir o exacerbar el sesgo en la estimación de los rendimientos. Si bien en el trabajo aquí presentado los instrumentos parecen ser válidos, hay muchos ejemplos en la literatura donde los instrumentos que se han utilizado son débiles, en particular, los instrumentos basados en el *background* familiar.

La evidencia sobre los rendimientos privados es globalmente convincente. A pesar de algunas sutilezas que la estimación plantea, la educación tiene un efecto claramente positivo sobre los ingresos del individuo. Por otra parte, la magnitud del efecto parece elevada en términos comparativos a otro tipo de inversiones. Se podría estar tentado a concluir que la existencia de altos rendimientos implica que existen beneficios para la sociedad (rendimientos sociales) por encima de los rendimientos privados, por lo que está justificado que los contribuyentes subvencionen a la educación. Pero también podría suceder que el rendimiento medido reflejase un rendimiento marginal, en cuyo caso deberíamos preguntarnos por qué este grupo marginal es tan rentable. Los diferentes enfoques adoptados en la literatura para obtener los rendimientos educativos a nivel macroeconómico constituyen la contribución central del trabajo de Sianesi y Van Reenen (2002).

## ■ AGRADECIMIENTOS

Los tres autores están afiliados al Instituto para el Estudio del Trabajo (IZA) en Bonn. Harmon y Walker agradecen el apoyo del Instituto para el Estudio del Cambio Social (CICS) en la UCD, por facilitar esta colaboración como parte del programa CICS en Evaluación de Políticas. Harmon es también investigador del University College de Londres, con el apoyo de la Fundación Nuffield bajo su nuevo régimen de becas de desarrollo profesional, a la que agradece su financiación. Agradecen al Departamento Británico de Educación y Trabajo y a Atlantic Philanthropies (Irlanda) por la financiación de esta actividad de investigación, y a la Comisión de la UE por la financiación de la red PURE, que realizó parte de la evidencia empírica presentada en este trabajo. Se agradece a Arnaud Chevalier, Denny Kevin, Conlon Gavan, Sianesi Barbara y John Van Reenen por sus comentarios sobre las versiones preliminares de este trabajo, y a nuestros colegas en el grupo PURE, en particular, a Rita Asplund y Pedro Pereira, por su cooperación. Los autores asumen cualquier responsabilidad sobre el contenido de este estudio.

## ■ BIBLIOGRAFÍA

- ANGRIST, J. y KRUEGER, A. B. (1991), "Does Compulsory School Attendance Affect Schooling and Earnings", *Quarterly Journal of Economics*, 106, 979-1014.
- (1992), "Estimating the Payoff to Schooling Using the Vietnam-Era Draft Lottery", *NBER Working Paper* no. 4067.
- ANGRIST, J. y NEWKEY, W. (1991), "Over-Identification Tests in Earnings Functions with Fixed Effects", *Journal of Business and Economic Statistics*, 9, 317-323.

- ASHENFELTER, O. y KRUEGER, A. (1994), "Estimates of the Economic Return to Schooling for a New Sample of Twins", *American Economic Review*, 84, 1157-1173.
- ASHENFELTER, O. y ROUSE, C. (1998), "Income, Schooling, and Ability: Evidence from a New Sample of Identical Twins", *Quarterly Journal of Economics*, 113, 253-284.
- ASHENFELTER, O. y ZIMMERMAN, D. (1997), "Estimates of the Return to Schooling from Sibling Data: Fathers, Sons and Brothers", *Review of Economics and Statistics*, 79, 1-9.
- ASHENFELTER, O.; HARMON, C. y OOSTERBEEK, H. (1999), "A Review of Estimates of the Schooling/Earnings Relationship with a Test for Publication Bias", *Labour Economics*, 6, 453-470.
- BECKER, G. (1964), *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education*, New York, Columbia University Press.
- BJORKLUND, A. y KJELLSTROM, C. (2000), "Estimating the Return to Investments in Education: How Useful is the Standard Mincer Equation?", *Economics of Education Review*, forthcoming.
- BLUNDELL, R.; DEARDEN, L. y MEGHIR, C. (1996), *The determinants and effects of work related training in Britain*, London, Institute for Fiscal Studies.
- BLUNDELL, R.; DEARDEN, L. y SIANESI, B. (2001), "Estimating the Returns to Education: Models, Methods and Results", *Journal of the Royal Statistical Society*, forthcoming.
- BONJOUR, D.; HASKEL, J. y HAWKES, D. (2000), "Estimating Returns to Education Using a New Sample of Twins", mimeo, Queen Mary and Westfield College.
- BORJAS, G. (1999), "The Economic Analysis of Immigration", en O. ASHENFELTER y D. CARD (eds.), *Handbook of Labor Economics*, Amsterdam and New York: North Holland.
- BOUND, J.; JAEGER, D. y BAKER, R. (1995), "Problems with Instrumental Variables Estimation when the Correlation Between the Instruments and the Endogenous Explanatory Variables is Weak", *Journal of the American Statistical Association*, 90, 443-450.
- BOUND, J. y JAEGER, D. (1996), "On the Validity of Season Of Birth as an Instrument in Wage Equations: A Comment on Angrist and Krueger's 'Does Compulsory School Attendance Affect Schooling and Earnings?'" , *NBER Working Paper no. 5835*.
- BOUND, J. y SOLON, G. (1999), "Double Trouble: On the Value of Twins-Based Estimation of the Return to Schooling", *Economics of Education Review*, 18, 481-500.
- BROWN, S. y SESSIONS, J. (1998), "Education, Employment Status and Earnings: A Comparative Test of the Strong Screening Hypothesis", *Scottish Journal of Political Economy*, 45, 586-591.
- BUCHINSKY, M. (1994), "Changes in the U.S. Wage Structure 1963-1987: Application of Quantile Regression", *Econometrica*, 62, 405-458.
- BUTCHER, K. y CASE, A. (1994), "The Effect of Sibling Composition on Women's Education and Earnings", *Quarterly Journal of Economics*, 109, 531-563.
- CARD, D. (1995), "Earnings, Schooling, and Ability Revisited", in S. Polacheck (ed.), *Research in Labor Economics*, Vol. 14. Greenwich: JAI Press.
- (1999), "Education and Earnings", en O. ASHENFELTER y D. CARD (eds.), *Handbook of Labor Economics*, Amsterdam and New York: North Holland.
- CARD, D. y KRUEGER, A. (1996), "Labor Market Effects of School Quality: Theory and Evidence", en G. BURTLESS (ed.), *Does Money Matter?* Washington, D.C.: Brookings Institution Press.

- CAWLEY, J.; HECKMAN, J. y VYTLACIL, E. (1998), "Cognitive Ability and the Rising Return to Education", *NBER Working Paper Series* no. 6388.
- CHEVALIER, A. (2000), *Graduate Overeducation in the UK*, mimeo, LSE Centre for the Economics of Education.
- CHEVALIER, A. y LYDON, R. (2001), *Estimates of the Effect of Education on Job Satisfaction*, mimeo, University of Warwick.
- DEARDEN, L. (1995), *Education, Training and Earnings*. PhD Thesis, University College London.
- (1998), *Ability, Families, Education and Earnings in Britain*. Institute for Fiscal Studies Working Paper no. W98/14.
- DENNY, K. y COLM, H. (2000), "Educational Policy Reform and the Return to Schooling from Instrumental Variables", IFS working paper no. W00/06.
- DESCHENES, O. (2001), *Unobserved Ability, Comparative Advantage and the Rising Return to Education in the United States: A Cohort Based Approach*. Mimeo, Princeton University.
- DOLTON, P. y VIGNOLES, A. (2000), "The Incidence and Effects of Overeducation in the UK Graduate Labour Market", *Economics of Education Review*, 19, 179-198.
- DUFLO, E. (1999), *Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment*. Mimeo, MIT.
- EVANS, W. y MONTGOMERY, E. (1994), "Education and Health: Where There's Smoke There's an Instrument", *NBER Working Paper*: 4949.
- GRILICHES, Z. (1977), "Estimating the Returns to Schooling: Some Econometric Problems", *Econometrica*, 45, 1-22.
- (1979), "Sibling Models and Data in Economics: Beginnings of a Survey", *Journal of Political Economy*, 87, S37-S64.
- GROOT, W. y MAASSEN VAN DEN BRINK, H. (2000), "Overeducation in the Labor Market: A Meta Analysis", *Economic of Education Review*, 19, 149-158.
- HAMERMESH, D. y BIDDLE, J. (1998), "Beauty, Productivity and Discrimination: Lawyers' Looks and Lucre", *Journal of Labor Economics*, 16, 172-201.
- HARMON, C. y WALKER, I. (1995), "Estimates of the Economic Return to Schooling for the United Kingdom", *American Economic Review*, 85, 1278-1286.
- (1999), "The Marginal and Average Return to Schooling in the UK", *European Economic Review*, 43, 879-887.
- (2000), "Returns to the Quantity and Quality of Education: Evidence for Men in England and Wales", *Economica*, 67, 19-35.
- HARMON, C., WALKER, I. y WESTERGAARD-NIELSEN, N. (2001), "Introduction", en C. HARMON, I. WALKER y N. WESTERGAARD-NIELSEN (eds.), *Education and Earnings in Europe*. Aldershot: Edward Elgar.
- HECKMAN, J. (1990), "Varieties of Selection Bias", *American Economic Review (Papers and Proceedings)*, 80, 313-328.
- HECKMAN, J.; LOCHNER, L. y TABER, C. (1999), "Human Capital Formation and General Equilibrium Effects: A Study of Tax and Tuition Policy", *Fiscal Studies*, 20, 25-40.
- HECKMAN, J.; LOCHNER, L. y TODD, P. (2001), *Fifty Years of Mincer Earnings Regressions*. Mimeo, University of Chicago.
- HECKMAN, J. y POLACHEK, S. (1974), "Empirical Evidence on the Functional Form of the Earnings-Schooling Relationship", *Journal of the American Statistical Association*, 69, 350-354.

- HUNGERFORD, T. y SOLON, G. (1987), "Sheepskin Effects in the Return to Education", *Review of Economics and Statistics*, 69, 175-177.
- ICHINO, A. y WINTER-EBMER, R. (2000), *The Long-Run Educational Cost of World War Two*. Mimeo, EUI Florence.
- IMBENS, G. W. y ANGRIST, J. D. (1994), "Identification and Estimation of Local Average Treatment Effects", *Econometrica*, 62, 467-475.
- ISACSSON, G. (1999), "Estimates of the Return to Schooling in Sweden From a Large Sample Of Twins", *Labour Economics*, 6(4), 471-489.
- KRUEGER, A. (2000), *Economic Considerations and Class Size*. Mimeo, Princeton University Industrial Relations Section.
- KRUEGER, A. y LINDAHL, M. (1999), "Education for Growth in Sweden and the World". *NBER working paper no. 7190*.
- LANG, K. (1993), *Ability Bias, Discount Rate Bias and the Return to Education*. Mimeo, Boston University.
- MASON, G. (1996), "Graduate Utilisation and the Quality of Higher Education in the UK", *NIESR discussion paper no. 158*.
- MEGHIR, C. y PALME, M. (1999), "Assessing the Effect of Schooling on Earnings Using a Social Experiment", *IFS working paper no. W99/10*.
- MILLER P.; MULVEY, C. y MARTIN, N. (1995), "What do Twins Studies Reveal about the Economic Return to Education? A Comparison of Australian and US Findings", *American Economic Review*, 85, 586-599.
- MINCER, J. (1974), *Schooling, Experience and Earnings*. New York: Columbia University Press.
- MURPHY, K. y WELCH, F. (1990), "Empirical Age-Earnings Profiles", *Journal of Labor Economics*, 8, 202-229.
- OECD (2001), *Education at a Glance*. Paris: OECD.
- PARK, J. H. (1996), "Measuring Education Over Time", *Economics Letters*, 60, 425-428.
- PEREIRA, P. y SILVA-MARTINS, P. (2002), "Is There a Return-Risk Link in Education?", *Economics Letters*, 75, (2), 31-37.
- PHILLIPSON, T. (2000), "External Treatment Effects and Program Implementation Bias", *NBER Technical Working Paper No. 250*.
- PONS, E. y GONZALO, M. T. (2001), *Returns to Schooling in Spain-How Reliable Are IV Estimates?*, mimeo, Queen Mary and Westfield College.
- PSACHAROPOULOS, G. (1983), "Education and Private Versus Public Sector Pay", *Labour and Society*, 8, 123-134.
- (1994), "Returns to Investment in Education: A Global Update", *World Development*, 22, 1325-1343.
- PSACHAROPOULOS, G. y LAYARD, R. (1979), "Human Capital and Earnings: British Evidence and a Critique", *The Review of Economic Studies*, 46, 485-503.
- RILEY, J. G. (2001), "Silver Signal-Twenty Five Years of Screening and Signalling", *Journal of Economic Literature*, 39, 432-478.
- RISCHALL, I. (1999), *The Roles of Education, Skill and Parental Income in Determining Wages*. Mimeo, McMaster University.
- ROUSE, C.E. (1999), "Further Estimates of the Economic Return to Schooling from a New Sample of Twins", *Economics of Education Review*, 18, 149-157.
- RUBIN, D. (1974), "Characterizing the Estimation of Parameters in Incomplete-Data Problems", *Journal of the American Statistical Association*, 69, 467-474.

- SIANESI, B. y VAN REENEN, J. (2002), "The Returns to Education-A Review of the Macroeconomic Literature", *Journal of Economic Survey*, this issue.
- SPENCE, M. (1973), "Job Market Signalling", *Quarterly Journal Of Economics*, 87, 355-373.
- TROSTEL, P.; WALKER, I. y WOOLEY, P. (2002), "Estimates of the Economic Return to Schooling in 28 Countries", *Labour Economics*, forthcoming.
- UUSILATO, R. (1999), "Returns to Education in Finland", *Labour Economics*, 6, 569-580.
- WILLIS, R. (1986), "Wage Determinants: A Survey and Reinterpretation of Human Capital Earnings Functions", en: O. ASHENFELTER y R. LAYARD (eds.), *Handbook of Labor Economics*. Amsterdam and New York: North Holland.

## ■ NOTAS

<sup>I</sup> La especificación de Mincer se ha ampliado para abordar cuestiones como la discriminación, la eficacia de los programas de formación (Blundell *et al.*, 1996), calidad de las escuelas (Card y Krueger, 1996), el rendimiento de las competencias lingüísticas (Borjas, 1999), e incluso el rendimiento de la "belleza" (Hammermesh y Biddle (1998)).

<sup>II</sup> Según esto, el modelo supone que la escolarización temprana tiene una rentabilidad mayor que la escolaridad en los últimos años de vida, ya que hay más períodos para recuperar los costes.

<sup>III</sup> En el modelo de Griliches hay una extensión sutil y frecuentemente ignorada pero que Card (1995) ha resalta-do. Puede existir una relación negativa entre la escolaridad óptima y el término de perturbación en la ecuación de salarios si se supone la existencia de un efecto no medido (lo podríamos denominar energía o motivación) que incrementa los ingresos y, por tanto, la renta dejada de percibir mientras se permanece en el proceso formativo, pero es independiente de los restantes costes de la escolarización.

<sup>IV</sup> Que la escolaridad y la habilidad sean factores complementarios en la producción de capital humano depende parcialmente del sistema escolar. Un sistema educativo que destine muchos recursos a complementar las deficiencias de los menos capacitados mostrará un distinto grado de complementariedad que otro sistema que oriente sus recursos hacia los individuos más capaces.

<sup>V</sup> Sin embargo, Cawley, Heckman y Vytlačil (1998) demuestran que la habilidad cognitiva y la escolarización medidas están tan altamente correlacionadas de suerte que no se puede separar sus efectos sin imponer restricciones que son rechazadas por los datos.

<sup>VI</sup> De hecho, para los datos de Gran Bretaña e Irlanda del Norte la información sobre la escolarización está limitada a los 18 años, por lo que es probable que el rendimiento estimado esté sesgado al alza.

<sup>VII</sup> Austria, Holanda, Grecia, España e Italia, usaron los salarios netos. Más detalles están disponibles en Harmon, Walker y Westergard-Nielsen (2001). Una alternativa es utilizar el Panel de Hogares de la Unión Europea (PHOGUE) del Eurostat. La ventaja del PHOGUE es, obviamente, que cada variable se ha especificado de la misma manera, independientemente del país. La desventaja, sin embargo, es que los datos del PHOGUE son inferiores respecto a la mayoría de bases de datos empleadas en este estudio, en términos de fiabilidad y número de observaciones.

<sup>VIII</sup> Si la ecuación de salarios es  $\log w_i = X_i\beta + \gamma S_i + \delta X_i + \gamma X_i^2 + u_i$ , entonces el ajuste implica restar  $\delta - 2\gamma(A - S)$ . Para un valor promedio de  $A-S$  alrededor de 20, el ajuste sería modesto, del orden del 2%.

<sup>IX</sup> El ajuste sugerido en la nota anterior, indica que las estimaciones del efecto de un año de educación controlando por la edad, es incluso menor que estas pequeñas diferencias que muestran los datos sin procesar.

<sup>X</sup> No se incluyeron controles por ocupación. Por lo general, controlar por ocupación reduce la estimación de los rendimientos educativos, ya que la estimación queda supeditada a la ocupación. En menor o mayor medida los rendimientos de la educación se deben a la posibilidad de alcanzar una mejor posición ocupacional, lo que es una consecuencia directa del nivel educativo alcanzado, en lugar de deberse al efecto que la educación tiene sobre los salarios controlando por ocupación.

<sup>xI</sup> Téngase en cuenta que el gráfico 3.3 simplemente agrupa los datos de salarios y escolaridad para las cualificaciones más altas, y por lo tanto, no se somete al control de otras diferencias entre los distintos grupos, como la edad. Dado que la edad se correlaciona positivamente con los salarios y negativamente con la educación, es probable que esta omisión cause un sesgo al alza en la estimación de los rendimientos educativos por mínimos cuadrados.

<sup>xII</sup> Krueger y Lindahl (1999) presentan cifras comparables para EE.UU., Suecia y Alemania. Dearden (1998). Sin embargo, demuestra que las titulaciones conseguidas pueden ser importantes en los sistemas educativos que se basan en gran medida en las cualificaciones, como en el caso del Reino Unido.

<sup>xIII</sup> Sobre esta parte de la distribución salarial pueden operar efectos derivados de la recesión de finales de los setenta.

<sup>xIV</sup> Deschenes (2001) utilizando datos de CPS para los Estados Unidos concluye que la rentabilidad de la educación aumentó en un 50% entre 1979 y 2000, pero que el crecimiento después de analizar los efectos de habilidad diferencial y sesgo de heterogeneidad entre las cohortes fue de aproximadamente 30%. No obstante, el efecto dominante es el problema de heterogeneidad (la correlación entre el rendimiento de la educación y el nivel educativo del individuo), con menores efectos que provienen de un aumento en el rendimiento de la habilidad inobservada, o de reforzamiento en la correspondencia habilidad-educación entre cohortes.

<sup>xV</sup> Sin embargo, hay que señalar que estas estimaciones no son independientes. Por ejemplo, múltiples estimaciones de los rendimientos educativos pueden ser obtenidas de un único estudio dentro de un país. Véase Krueger (2000) para una discusión sobre la implicación de esto en un metanálisis de los efectos tamaño de clase.

<sup>xVI</sup> La implicación de que se de correlación entre educación y salarios y que esta sea debida a señalización en lugar de a capital humano es que la rentabilidad social de la educación será menor que la privada. Añadir un año extra de educación a todos los individuos no tiene ningún efecto sobre la productividad y no mejora la capacidad de elección de los empleadores. En consecuencia, aumentar la educación no tiene efecto alguno sobre el PIB. Es por tanto importante diferenciar entre señalización y capital humano porque la misma rentabilidad privada de la educación puede tener muy distintas implicaciones para la rentabilidad social. Ver Sianesi y Van Reenen (2002).

<sup>xVII</sup> En este ejemplo, la fuente de correlación entre  $s$ , y  $\epsilon$ , es que una variable explicativa relevante ha sido omitida. Otras fuentes de dicha correlación podrían ser errores de medición en  $s$ , y sesgo de selección.

<sup>xVIII</sup> El estudio de Angrist y Krueger ha sido criticado por Bound y Jaeger (1996) que argumentan, que el trimestre de nacimiento puede tener un impacto en los ingresos que no opere a través de su efecto en la escolarización.

<sup>xIX</sup> En términos generales, los experimentos naturales ignoran la posibilidad de que el tratamiento puede tener un impacto en las personas que no sean del grupo tratamiento. Si, por ejemplo, la edad de escolarización obligatoria se eleva, los que abandonan la escuela justo antes del aumento pertenecerán a un grupo de personas poco educadas que no tienen las competencias de las cohortes más jóvenes, con el mismo bajo nivel de escolaridad. Esto puede aumentar los salarios de este grupo, y dar lugar a un sesgo a la baja en el rendimiento estimado de un año de escolaridad. No obstante, esto puede no ser un tema importante, en la medida en que el efecto tratamiento tienda a afectar al flujo, por ejemplo, a través de alguna reforma, mientras que el *stock* de población tienda a ser elevado. Véase Philipson (2000), para mayor discusión sobre este argumento.

<sup>XX</sup> El supuesto implícito en este procedimiento es que el rendimiento de los años de escolaridad continua, es el mismo que el rendimiento educativo cuando se reanuda la escolarización tras una interrupción, lo que podría no ser realista.





4

## **CAPITAL HUMANO Y RENDIMIENTOS DE LA EDUCACIÓN EN ESPAÑA**

José L. Raymond  
José L. Roig  
Gustavo A. García  
Lina M. Gómez



## 4. CAPITAL HUMANO Y RENDIMIENTOS DE LA EDUCACIÓN EN ESPAÑA

### ■ 4.1. INTRODUCCIÓN

La educación cumple un papel esencial en la economía. La acumulación de capital humano y el correcto aprovechamiento de este por parte del aparato productivo, pueden asegurar el crecimiento económico y el desarrollo de las condiciones de vida para la población. Este factor le concede sin duda un interés público a la educación. Sin embargo, el interés privado que pueden tener los individuos respecto a educarse o no, está directamente condicionado por la valoración diferencial que el mercado le concede a los más educados, frente a los no educados.

La cuantificación de los retornos educativos, refleja en cuanto se incrementa porcentualmente el salario del individuo cuando este adquiere un año de educación adicional. A mayor rentabilidad de la educación, los individuos se verán incentivados a acumular capital humano como modo de diferenciarse en la sociedad.

La estimación de los retornos a la educación está ampliamente extendida siguiendo la metodología propuesta por Mincer (1974), o aplicando ligeras variaciones. A pesar de la aceptación generalizada de este procedimiento, tampoco está exento de críticas en cuanto a su condicionamiento respecto a la información disponible y la fiabilidad de la misma.

En este documento, proponemos hacer un repaso sobre los retornos educativos y su relación con la dotación de capital humano, evaluando diversos aspectos como las diferencias por sexo, la evolución temporal, su relación con el desempleo, la corrección del sesgo de habilidad, los factores condicionantes de su evolución para España y la presencia de externalidades del capital humano.

Para el desarrollo de esta investigación, se empleó la información de la Encuesta de Estructura Salarial (EES), para los periodos 1995, 2002 y 2006, lo que permitió efectuar una comparación completa y homogénea en el tiempo. Esta valiosa fuente de información se complementó con la que proporciona la Encuesta de la Población Activa (EPA) para los mismos años, y la Encuesta de Condiciones de Vida (ECV) para 2005. La metodología empleada se describe en cada apartado.

De este estudio se concluye que España durante el periodo 1995 a 2006, ha sido objeto de una caída de los retornos a la educación en casi 2 puntos porcentuales. Los retornos estimados en este estudio pasaron aproximadamente de un 8,7% a 6,9%, durante el periodo de análisis. Adicionalmente, se observa una aparente reducción en el salario medio de la población, que esta influenciada fundamentalmente por el cambio en las características de la población<sup>1</sup>. Si se compara el mismo individuo típico en el tiempo, se obtiene que bajo la hipótesis *ceteris paribus*, la caída de salarios no se produce.

Se evidencian diferencias importantes en los retornos estimados por sexo y por región, según nivel educativo. La composición del capital humano en cada caso, juega un papel muy importante en el comportamiento de la rentabilidad de la educación. En general, el mercado laboral español ha sido objeto de una transformación demográfica que aumentó la participación laboral de la mujer y el nivel educativo promedio de la población (siendo este incluso mayor para las mujeres). Estos aspectos condicionados a los requerimientos del sistema productivo de cada región, y al mayor o menor ajuste del mercado laboral son factores que determinan claramente las diferencias de retornos encontradas. Contra el cambio demográfico, se propone una recomposición de la estructura productiva hacia mejoramientos en la composición tecnológica que requiera mano de obra más cualificada, como una forma de contrarrestar la tendencia a la baja de los retornos educativos.

Esto tiene especial interés teniendo en cuenta que en el capítulo 10 de este estudio se confirma la presencia de externalidades positivas de la acumulación de capital humano cualificado, que no sólo afecta la productividad de la mano de obra cualificada, sino también sobre la productividad de la mano de obra no cualificada. Este resultado confirma la importancia que tiene la acumulación de capital humano para la sociedad en su conjunto, pero para que ello efectivamente se dé debe coexistir el incentivo privado a educarse. Es decir, los individuos decidirán acumular capital humano, sólo si los más educados pueden ubicarse de manera efectiva dentro del mercado laboral y logran además obtener una retribución importante en términos de salarios por tomar dicha decisión<sup>2</sup>; y esto únicamente sucederá cuando el sistema productivo se desarrolle tecnológicamente y requiera en mayor proporción mano de obra cualificada.

Finalmente, la corrección del sesgo de habilidad empleando dos metodologías (pseudopanel y variables instrumentales) y bases de datos (EES y ECV) distintas, presenta una corrección al alza de los retornos educativos, confirmando los resultados presentados por otros estudios con datos españoles (Pons y Gonzalo, 2002; Arrazola *et al.*, 2003; Arrazola y Hevia, 2006; Arrazola y Hevia, 2008), si bien tal resultado es contrario a lo que la intuición sobre la corrección del sesgo de endogeneidad sugiere.

Este documento ofrece una breve revisión de la literatura y una descripción de las bases de datos empleadas, el análisis descriptivo de las características del capital humano

<sup>1</sup> La más importante quizá la significativa reducción de los años de antigüedad promedio.

<sup>2</sup> Lo cual además implica que la decisión es también eficiente en términos económicos.

según regiones y sectores económicos, y la estimación de los retornos educativos y su relación con el desempleo, identificando las diferencias por sexo y nivel educativo en el tiempo. Se presenta, además, la estimación de los retornos educativos por dos metodologías diferentes para corregir el denominado sesgo de habilidad. Finalmente, se aporta evidencia sobre los factores determinantes del comportamiento de los salarios y sobre la presencia de externalidades del capital humano en el mercado laboral español. Conclusiones generales, bibliografía y anexos complementan el estudio.

#### ■ 4.2. CAPITAL HUMANO Y RENDIMIENTOS EDUCATIVOS EN ESPAÑA. ALGUNOS ESTUDIOS DE SU EVOLUCIÓN Y DISTRIBUCIÓN GEOGRÁFICA

El acervo de artículos y estudios sobre el capital humano y sus rendimientos en España es amplio. En esta sección más que ser exhaustivos en el repaso de la literatura se van a referenciar dos trabajos que han contribuido a la elaboración del nuestro y van en la misma línea que se pretende desarrollar. Dichos trabajos son: Serrano y Pastor (2005) y Pastor, Raymond, Roig y Serrano (2006). Estos estudios se enmarcan dentro de los proyectos de investigaciones sobre el capital humano desarrollados por Bancaja y el IIE.

El primer trabajo aborda en profundidad la distribución geográfica del capital humano en España a partir de los datos del Censo de 2001. Entre los resultados se encuentra que en el total de España el nivel de capital humano, medido como los años medios de estudio, fueron de 7,6 años. Por comunidad autónoma se encuentra que el País Vasco y Madrid son las que presentan los mayores niveles, con 8,7 y 8,6 años medios de estudio respectivamente. En cuanto a las comunidades autónomas con menores años de estudios en su población, se encuentran Extremadura y Castilla-La Mancha, con 6,5 años medios cada una. La distribución geográfica muestra que espacialmente se configuran dos áreas: el nordeste de la península con mayores niveles de capital humano y el suroeste con los menores niveles.

Serrano y Pastor (2005) además de analizar el capital humano de toda la población española, estudian el comportamiento del capital humano productivo, es decir, el capital humano utilizado para producir bienes y servicios, medido a partir de los años medios de educación de los ocupados. Los resultados de los autores muestran que el anterior indicador para el conjunto de España en el 2001 toma el valor de 10,26 años, el cual es 35% mayor a los niveles educativos para el conjunto de la población. Éste comportamiento es generalizado en todas las regiones y se nota una clara desigualdad en el capital humano productivo per cápita. Mientras que el nordeste junto a Madrid tienen una diferencia de 11 años de estudio entre la población total y la ocupada, en el extremo opuesto se encuentran las comunidades de sur con una diferencia de alrededor de 9 años.

El trabajo de Pastor, Raymond, Roig y Serrano (2006), por su parte, se centra en el análisis del aprovechamiento del capital humano y su rendimiento en España, y se hace una comparación con otros países del entorno. Dentro de los principales resultados sobre dota-

ción de capital humano se ha encontrado que España, a pesar de tener mejoras considerables en los niveles educativos en el periodo 1964-2005, en la comparación a nivel internacional muestra que se encuentra entre los países con menores dotaciones de capital humano de la OCDE. En cuanto al uso del capital humano los resultados muestran que España ha aumentado el grado de utilización de su capital humano potencial, pero es menor al uso que le dan la mayoría de los países europeos desarrollados.

Pasando al análisis de los rendimientos de la educación, las estimaciones de los autores muestran que para 2001 España presenta un rendimiento promedio alrededor del 8,5%, el cual, comparado con el rendimiento de 11,5% de Irlanda, que fue el más alto, puede considerarse elevado. Por su parte las estimaciones a nivel nacional mostraron que en 1995 los rendimientos de la educación en España se ubicaban en 8,66%, mientras que en 2002 fueron de 7,86%. Se produce, en consecuencia, una caída de 0,8 puntos porcentuales. A nivel de comunidades autónomas para 2002 se observó que Madrid presenta la tasa de rendimiento más elevada, 9,2%. Cuando las estimaciones se controlan por la composición sectorial de las regiones, Madrid sigue liderando pero con una tasa más baja, 8,92%, seguida de Cataluña con un rendimiento de 7,52%. Entre las comunidades autónomas con menores rendimientos de la educación, se encontraron a Cantabria (5,53%), Murcia (5,5%), Castilla y León (5,49%), y Extremadura (5,49%).

#### ■ 4.3. BASE DE DATOS DISPONIBLES Y HOMOGENEIZACIÓN. ANÁLISIS DESCRIPTIVO

En este estudio se pretende hacer un análisis de los principales resultados sobre los retornos educativos en España, y para ello se empleará la información de la *Encuesta de Estructura Salarial* (EES) y de la *Encuesta de Condiciones de Vida* (ECV). La *Encuesta de Población Activa* (EPA), se empleará de manera complementaria para obtener información sobre el comportamiento de los individuos en el mercado laboral. Estas fuentes de información han sido elaboradas por el Instituto Nacional de Estadística de España (INE).

La EES y la ECV aportan relevante información sobre la población ocupada. No obstante, se diferencian dado que la EES se especializa en información sobre el puesto de trabajo y sólo incluye a los ocupados por cuenta ajena, mientras que la ECV es más rica en información sociodemográfica y es representativa del conjunto de la población en edad de trabajar. En cualquier caso, ambas encuestas aportan información sobre niveles educativos y salarios, variables fundamentales para el análisis de rendimientos educativos<sup>3</sup>.

La *Encuesta de Estructura Salarial* (EES), es una encuesta que investiga la distribución de los salarios en función de una gran variedad de variables como son el sexo, la ocupa-

<sup>3</sup> Hasta 2001 se contó con la información del Panel de Hogares de la Unión Europea (PHOGUE), que fue reemplazado por la ECV.

Tabla 3.1

## CARACTERÍSTICAS DE LOS ASALARIADOS SEGÚN LA MUESTRA

VARIABLES	1995	2002	2006
<b>Medias (Desviación estándar)</b>			
Edad	38,3 (10,9)	37,1 (10,8)	37,7 (10,7)
Años de Escolaridad	8,6 (3,7)	9,0 (3,9)	9,1 (4,1)
Años de Experiencia (potencial)	23,7 (12,2)	22,1 (12,0)	22,6 (11,9)
Años de Antigüedad	10,2 (9,8)	7,0 (9,3)	6,1 (8,7)
Salario mensual real (euros)	1839,1 (1352,0)	1786,3 (1461,4)	1705,1 (1269,0)
Salario mensual real/hora (euros)	11,4 (8,4)	11,3 (8,9)	10,8 (7,8)
<b>Frecuencias (%)</b>			
Sexo			
Hombre	76,3	68,8	65,2
Mujer	23,7	31,2	34,8
Nacionalidad			
Español	ND	96,7	92,2
Extranjero	ND	3,3	7,8
<b>Tamaño Muestral</b>	140.297	149.405	112.119

Fuente: Cálculos propios a partir de las EES.

ción, la rama de actividad, la antigüedad, o el tamaño de la empresa, entre otros. Tiene periodicidad cuatrienal a partir de 2002, pero se dispone de información para los años 1995, 2002 y 2006. Se realiza en todos los Estados miembros de la Unión Europea.

La población objetivo, la comprenden los trabajadores por cuenta ajena que prestan sus servicios en centros de cotización, y que hayan estado de alta en la seguridad social durante todo el mes de octubre. A partir de 2006, la muestra se amplió para ser representativa de empresas con 1 a 9 empleados. Se estudian las actividades económicas encuadradas en tres grandes sectores: industria, construcción y servicios, comprendidas en las secciones de la C a K y de la M a la O de la CNAE-93<sup>4</sup>, quedando excluidos las actividades agrícolas, ganaderas y pesqueras, la Administración Pública, Defensa y Seguridad Social obligatoria, personal doméstico y organismos extraterritoriales.

Esta encuesta aporta amplia información sobre las características de los individuos y las empresas, es representativa del territorio nacional, y para 18 comunidades autónomas. Se dispone de desagregación sectorial a dos dígitos. Las tablas 3.1 y 3.2 contienen información estadística proporcionada por la muestra de la EES.

<sup>4</sup> En 1995, la encuesta no incluía los las secciones M, N y O, correspondientes a educación; actividades sanitarias y veterinarias, servicios sociales; y otras actividades sociales y de servicios prestados a la comunidad y servicios personales, respectivamente.

Tabla 3.2

## CARACTERÍSTICAS DE LOS EMPLEOS SEGÚN LA MUESTRA

VARIABLES	1995	2002	2006
	FRECUENCIA (%)		
<b>Tipo de jornada</b>			
Tiempo completo	96,3	91,0	86,6
Tiempo parcial	3,8	9,0	13,4
<b>Duración del contrato</b>			
Duración indefinida	73,3	74,0	72,9
Duración determinada	26,7	26,0	27,1
<b>Convenio Salarial</b>			
De sector	34,9	35,8	39,4
Interprovincial, provincial o comarcal	41,8	47,5	50,8
De empresa, centro de trabajo u otro	23,3	16,7	9,8
<b>Tamaño de la Empresa*</b>			
De 10 a 49	48,0	40,9	44,8
50 a 199	29,7	30,3	26,4
200 y más trabajadores	22,3	28,9	28,8
<b>Actividad Económica**</b>			
<b>Extracción prod. energéticos y no energéticos</b>			
Alimentación, bebidas y tabaco	7,9	6,1	5,4
Textil y de la confección	4,4	3,2	2,1
Cuero y del calzado	1,9	1,3	0,8
Madera y del corcho	2,5	2,5	2,0
Papel, edición, artes gráficas y reproducción	4,1	3,4	2,9
Refino petróleo y combustibles nucleares	0,1	0,0	0,0
Industria química	4,6	3,2	2,6
Caucho y materias plásticas	2,5	2,5	2,1
Productos minerales no metálicos	4,9	3,7	3,9
Metalurgia y fab. productos metálicos	5,4	5,7	5,3
Maquinaria y equipo mecánico	3,5	3,3	2,0
Mat. y equipo eléctrico, electrónico y óptico	4,1	3,5	2,3
Material de transporte	4,1	3,3	2,4
Manufactureras diversas	3,6	2,9	2,9
Energía eléctrica, gas y agua	2,0	1,7	1,3
Construcción	7,4	9,5	12,3
Comercio y reparaciones	10,6	11,7	13,7
Hostelería	6,0	6,7	7,2
Transporte, almacenamiento y comunicación	6,3	6,7	6,4
Intermediación financiera	6,9	6,1	6,1
Act. Inmobiliarias y servicios empresariales	5,5	11,7	15,4
<b>Tamaño Muestral</b>	<b>140.297</b>	<b>149.405</b>	<b>112.119</b>

\* Excluidos de la muestra para 2006 las empresas con menos de 10 empleados. \*\* Excluidas de la muestra para 2002 y 2006: educación, actividades sanitarias y veterinarias, servicios sociales, y otras actividades sociales y de servicios prestados a la comunidad, y servicios personales.

Fuente: Cálculos propios a partir de la EES.

Estas tablas se han elaborado partiendo de la homogeneización de la muestra para los periodos considerados. Por ello incluye sólo trabajadores en empresas de 10 y más empleados, cuya actividad se desarrolla en los sectores C a la K de la CNAE-93. Para la población de asalariados correspondiente con esta muestra, en la tabla 3.1 se observa un aumento de los años de escolaridad promedio, frente a una reducción en los niveles de antigüedad en la empresa, un deterioro en los salarios mensual y por hora. De hecho, una posible explicación de que el salario mensual real por hora decrezca puede ser debido a que se reduce la antigüedad en el puesto de trabajo. La composición de la población ha cambiado en la última década, contando con una mayor proporción de mujeres y población extranjera.

En la tabla 3.2, vemos un aumento en la participación de los contratos a tiempo parcial, y de los convenios salariales colectivos de tipo regional. Las empresas de 10 a 50 empleados aportan en promedio un 45% del empleo para esta población. La composición según sector de actividad presenta para 2006 un aumento en la participación de los sectores construcción y servicios, este último liderado por comercio y reparaciones, y actividades inmobiliarias y servicios empresariales.

La *Encuesta de Población Activa (EPA)*, está orientada a dar datos de las principales categorías poblacionales en relación con el mercado de trabajo (ocupados, parados, activos, inactivos). Su periodicidad es trimestral, si bien la metodología de recolección de información es continua. Es representativa del conjunto nacional, y se puede desagregar a nivel de comunidades autónomas y provincias. La unidad de análisis son las viviendas y las personas.

Las tablas 3.3 y 3.4 recogen algunas estadísticas sobre la composición y características del mercado laboral según la EPA.

En la tabla 3.3, observamos el cambio en la distribución de la población española según actividad en los últimos años. En primer lugar, entre 1995 y 2002, se refleja la inversión de la pirámide poblacional en la reducción de más de 2 puntos porcentuales en la proporción de menores. No obstante, esta tendencia se frena un poco en 2006, posiblemente por efecto de una mayor participación de la población extranjera. Por otra parte, la proporción de ocupados y parados respecto a la población total, refleja el efecto del ciclo económico expansivo en la economía española en los últimos años. Es posible que este efecto incida sobre la disminución de inactivos, en el caso en que el ciclo expansivo incentive a los jóvenes a buscar empleo, frente a continuar educándose.

Por otra parte, respecto a la situación profesional de los ocupados, cabe resaltar que para el año 2006, al menos un 82% de la población eran ocupados por cuenta ajena (asalariados). Esta cifra aumentó desde 1995 con el ciclo expansivo, ganando participación a los trabajadores independientes (autónomos), entre otros.

En la tabla 3.4, se presentan algunas características de la fuerza de trabajo. En primer lugar, se confirma un aumento en los años de escolaridad promedio tanto de los asalariados, como de los activos en general. Las mujeres presentan un nivel educativo promedio mayor que los hombres para ambos grupos de población, aunque su participación fue tam-

Tabla 3.3

## COMPOSICIÓN DE LA POBLACIÓN ESPAÑOLA EN EL MERCADO LABORAL

VARIABLES	1995		2002		2006	
	N.º	%	N.º	%	N.º	%
<b>Composición de la población según actividad</b>						
Menores	6.908.697	17,9	6.478.779	15,7	6.838.136	15,6
Ocupados	12.593.008	32,5	16.825.375	40,7	19.895.631	45,3
Parados	3.869.478	9,6	2.211.803	5,4	1.765.041	4,0
Inactivos	15.493.797	40,0	15.808.858	38,3	15.404.095	35,1
<i>Población Total</i>	<i>38.702.877</i>	<i>100</i>	<i>41.324.814</i>	<i>100</i>	<i>43.902.903</i>	<i>100</i>
<b>Situación profesional respecto al empleo</b>						
Empleador	620.643	4,9	891.909	5,3	1.094.384	5,5
Trabajador Independiente	1.942.870	15,4	1.862.616	11,1	2.068.758	10,4
Asalariado	9.413.653	74,8	13.698.816	81,4	16.365.853	82,3
Ayuda Negocio Familiar	473.199	3,8	269.857	1,6	255.247	1,3
Otra situación	142.642	1,1	102.178	0,6	111.390	0,6
<i>Total Ocupados</i>	<i>12.593.008</i>	<i>100</i>	<i>16.825.375</i>	<i>100</i>	<i>19.895.631</i>	<i>100</i>
<b>Tamaño Muestral</b>	191.021		174.349		159.511	

Fuente: Cálculos propios a partir de la EPA.

Tabla 3.4

## CARACTERÍSTICAS DE LA FUERZA DEL TRABAJO\*

VARIABLES	1995	2002	2006
<b>Asalariados</b>			
Edad	37,2 (11,3)	37,3 (11,1)	37,7 (11,3)
Años de Escolaridad	9,1 (4,2)	10,0 (4,1)	10,4 (4,0)
Escolaridad Hombres	8,7 (4,1)	9,6 (4,0)	10,0 (3,9)
Escolaridad Mujeres	9,9 (4,3)	10,7 (4,1)	11,0 (4,0)
Proporción Mujeres (%)	35,2%	39,8%	42,4%
<b>Activos</b>			
Edad	36,9 (12,1)	37,9 (11,6)	38,4 (11,7)
Años de Escolaridad	8,6 (4,1)	9,7 (4,1)	10,2 (4,0)
Escolaridad Hombres	8,3 (4,0)	9,3 (4,0)	9,8 (3,9)
Escolaridad Mujeres	9,2 (4,3)	10,3 (4,2)	10,7 (4,0)
Proporción Mujeres (%)	37,8%	40,1%	41,7%

\*Medias/Entre () desviación estándar. Fuente: Cálculos propios a partir de la EPA.

bién significativamente menor (35% para 1995), a pesar de aumentar en los últimos años. Esto implica que la participación de las mujeres ha sido en gran medida un fenómeno incentivado por el acceso de ésta al sistema educativo.

Por último, la *Encuesta de Condiciones de Vida (ECV)*, surge como la adaptación del *Panel de Hogares de la Unión Europea (PHOGUE)* a las nuevas necesidades de información de organismos públicos y privados. El primer año con información para España fue 2004 y su última publicación corresponde a 2007. Su información corresponde a un panel longitudinal, con importante información sobre características sociodemográficas del individuo y de su hogar. Cada año esta encuesta desarrolla un módulo especial, que en 2005 se dedicó a la "Transmisión intergeneracional de la pobreza" que incluye información sobre el nivel educativo de los padres, y que pretende aprovecharse en esta investigación para el cálculo de los retornos controlando por presencia de endogeneidad. Este módulo se dirige a los mayores de 25 años, refiriendo las preguntas al período de adolescencia del individuo.

Abordar el problema de la endogeneidad consiste en estimar los retornos controlando por el posible efecto que sobre los retornos tienen el hecho de que los padres procedan de un entorno educado o no. La lógica del planteamiento es que los individuos más hábiles son aquellos que tienden a adquirir un mayor nivel educativo, y dado que la habilidad no es observable y que sus efectos quedan recogidos por la perturbación aleatoria, ello provocará correlación entre la variable "años de escolaridad" y perturbación aleatoria. No obstante, si los padres más educados transmiten a sus hijos una mayor tendencia a educarse pero no transmiten a sus hijos una mayor habilidad innata, la educación de los padres será un instrumento válido para estimar rendimientos educativos. Dejará, no obstante, de ser un instrumento válido si el componente de transmisión genética de habilidad entre padres e hijos es importante.

Es decir, si a mayor habilidad innata le corresponde un mayor nivel educativo, la situación deseable para que la educación de los padres sea un instrumento válido es que entre padres e hijos exista transmisión de la preferencia por educarse, pero que entre padres e hijos no exista transmisión de habilidades innatas. De hecho, disponer de una muestra de hijos adoptivos ofrecería un buen instrumento, en la medida en que los padres más educados transmitan a sus hijos adoptivos una mayor preferencia por educarse, pero, por definición, genéticamente no pueden transmitir habilidad innata. Las tablas 3.5 y 3.6, presentan algunas estadísticas sobre la relación existente entre educación entre padres e hijos.

En primer lugar, en la tabla 3.5 confirmamos que los hijos tienen un nivel educativo superior al de los padres. Esto claramente está influenciado por la transición demográfica que tuvo lugar entre ambas generaciones. Adicionalmente, estas variables presentan una correlación positiva de 0,43, lo cual implica que a mayor nivel educativo de los padres, mayores años educación presentan los hijos. Los padres podrían por tanto, incidir sobre las habilidades de los hijos para adquirir mayores niveles de educación.

En la tabla 3.6, contiene la matriz de transición educativa entre padres e hijos. Las frecuencias ubicadas sobre la diagonal, indican en que casos los hijos tienen al menos el

Tabla 3.5

## ESCOLARIDAD DE PADRES E HIJOS

ESTADÍSTICAS	AÑOS DE ESCOLARIDAD	
	HIJOS	PADRES
Media	10,5	6,0
Mediana	8	5
Máximo	17	17
Mínimo	2,5	2,5
Desviación Estándar	4,7	4,0
Correlación Padres e Hijos	0,43	
Número de Observaciones	15.204	15.204

Fuente: Cálculos propios a partir de la ECV 2005.

Tabla 3.6

## RELACIÓN DEL NIVEL EDUCATIVO ENTRE PADRES E HIJOS

AÑOS DE ESCOLARIDAD	FRECUENCIA (%)						% HIJOS	ACUM. % HIJOS
	PADRES							
	2,5	5	8	12	14	17		
<b>Hijos</b>								
2,5	1,2%	0,3%	0,0%	0,0%	0,0%	0,0%	<b>1,5%</b>	1,5%
5	10,8%	14,2%	0,3%	0,2%	0,0%	0,2%	<b>25,7%</b>	27,2%
8	6,9%	14,2%	1,7%	0,5%	0,0%	0,4%	<b>23,7%</b>	50,9%
12	3,0%	12,4%	1,7%	1,9%	0,1%	1,6%	<b>20,8%</b>	71,7%
14	0,3%	1,0%	0,1%	0,1%	0,1%	0,1%	<b>1,7%</b>	73,4%
17	1,9%	14,0%	2,1%	2,8%	0,2%	5,7%	<b>26,6%</b>	100,0%
<b>% Padres</b>	<b>24,2%</b>	<b>56,0%</b>	<b>5,9%</b>	<b>5,5%</b>	<b>0,4%</b>	<b>8,0%</b>	<b>100,0%</b>	
<i>% acum. Padres</i>	24,2%	80,2%	86,0%	91,5%	92,0%	100,0%		

AÑOS DE ESCOLARIDAD	NÚMERO DE OBSERVACIONES EN LA MUESTRA						N.º OBS. HIJOS
	PADRES						
	2,5	5	8	12	14	17	
<b>Hijos</b>							
2,5	189	39	4	0	1	1	<b>234</b>
5	1.643	2.156	40	36	2	25	<b>3.902</b>
8	1.048	2.158	251	77	4	63	<b>3.601</b>
12	453	1.889	264	286	19	249	<b>3.160</b>
14	44	153	20	12	9	19	<b>257</b>
17	295	2.121	314	426	30	864	<b>4.050</b>
<b>No. Obs. Padres</b>	<b>3.672</b>	<b>8.516</b>	<b>893</b>	<b>837</b>	<b>65</b>	<b>1.221</b>	<b>15.204</b>

Fuente: Cálculos propios a partir de la ECV 2005.

mismo nivel educativo que sus padres. Claramente, los casos se concentran en la parte inferior de esta diagonal, confirmando que son los hijos de los padres más educados los que presentan mayores niveles educativos.

Una primera exploración descriptiva de las fuentes a utilizar en este estudio, nos presenta la clara transición demográfica que ha sufrido España en los últimos años, coincidente con un envejecimiento de la población, un aumento en los niveles de escolaridad y en las tasas de participación de la mujer. Por otra parte, a pesar que el ciclo económico ha sido claramente expansivo, ha estado acompañado de una reducción de los niveles de salario mensual y por hora, aunque esta reducción puede estar influenciada por el cambio en la composición de la población asalariada en los últimos años. El objetivo final de este estudio es indagar que factores han podido influenciar sobre la evolución de los retornos a la educación.

#### ■ 4.4. ANÁLISIS DEL COMPORTAMIENTO DEL CAPITAL HUMANO POR REGIONES Y SECTORES

En esta sección a partir de los datos del IVIE se analiza cómo es la composición del capital humano productivo en las comunidades autónomas y en los sectores económicos, así como su evolución entre 1995 y 2006.

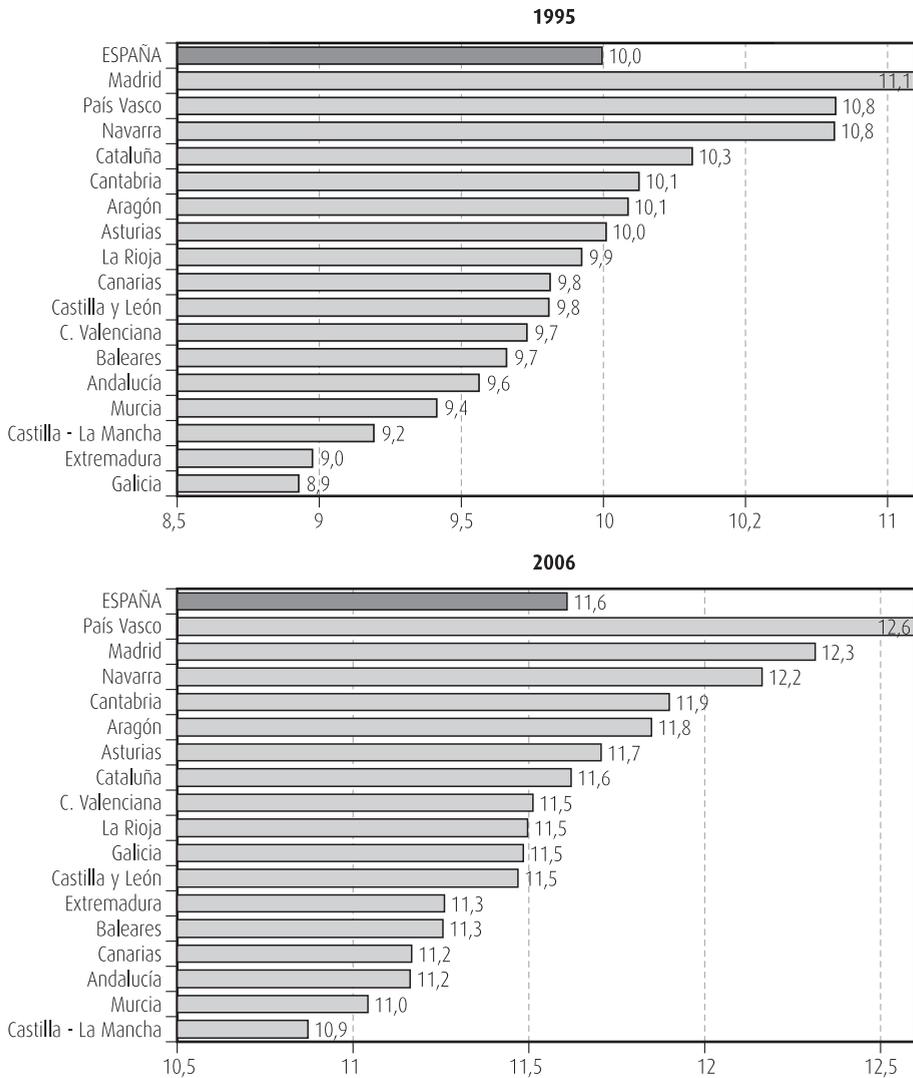
Un panorama regional del nivel de capital humano productivo, medido como los años medios de educación de los ocupados, se muestra en los gráficos 4.1 y 4.2. En el total de España, los años medios de estudio de los ocupados en 2006 fueron 11.6 años, casi dos años más en comparación con 1995. Este mismo comportamiento se observa a nivel de comunidad autónoma y se notan ciertos segmentos regionales asociados a las diferentes dotaciones en capital humano productivo. Por ejemplo, las regiones con mayores años medios de educación son, en su orden: Madrid, País Vasco y Navarra, con más de 12 años medios de educación en sus ocupados en 2006. En este primer grupo también se incluyen las comunidades de Cataluña, Aragón, Cantabria y Asturias, que se encuentran por encima del nivel medio nacional para los dos años en análisis.

Por su parte, las regiones que presentan cierta persistencia en tener menores años medios de educación son: Castilla-La Mancha, Murcia y Andalucía. Cabe resaltar que en 1995 Extremadura y Galicia constituían el grupo de comunidades autónomas con menores años medios de educación, pero para 2006 estas comunidades tienen un importante crecimiento de esta variable: pasan de 9 más de 11 años medios de educación.

Respecto a la situación geográfica de cada comunidad, se observan dos áreas geográficas diferenciadas en lo que se refiere a la distribución del capital humano productivo: el nordeste de la península con mayores niveles de capital humano y el suroeste con los menores niveles. Estos resultados son similares a los encontrados en los trabajos referenciados en la sección anterior.

Gráfico 4.1

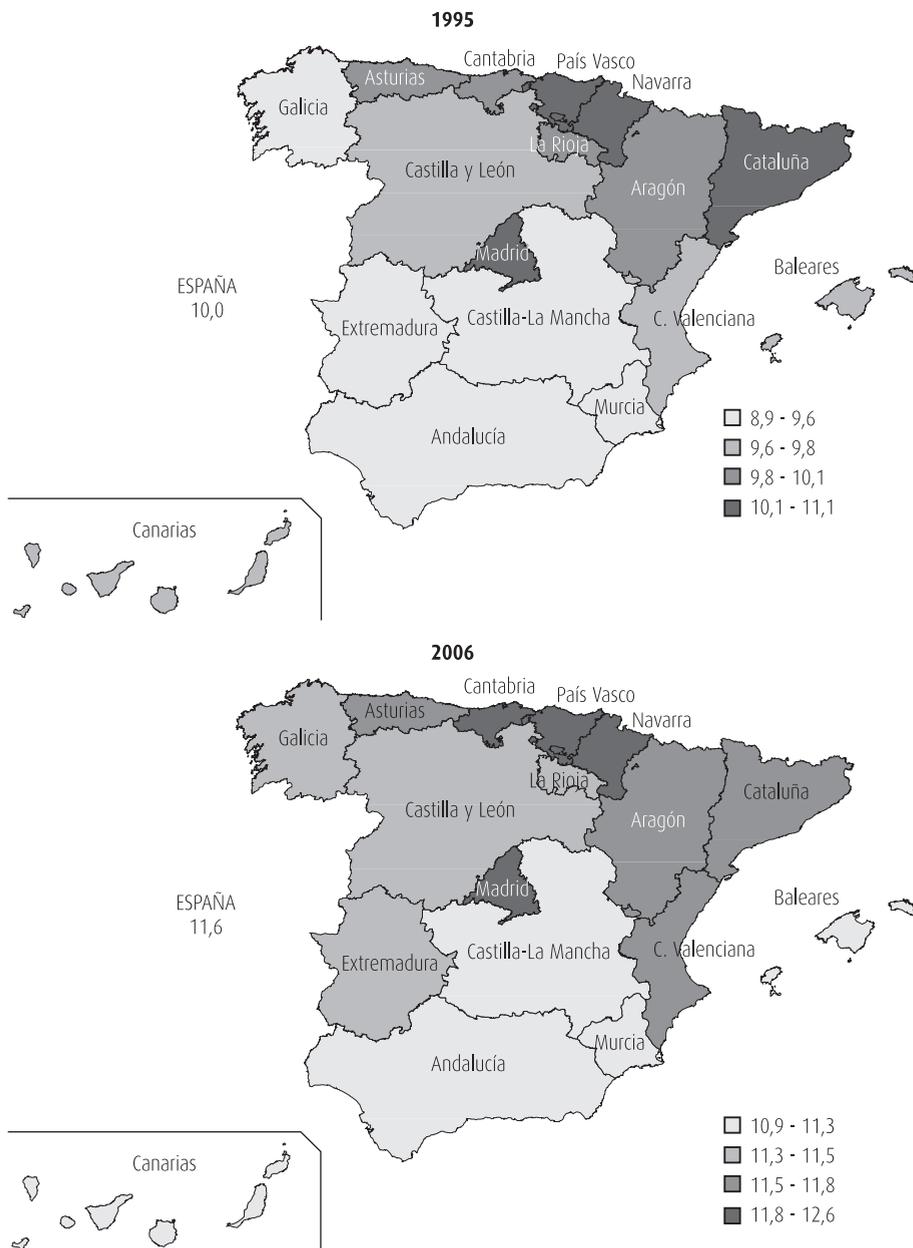
## AÑOS MEDIOS DE EDUCACIÓN DE LA POBLACIÓN OCUPADA POR CCAA



Fuente: Cálculos propios a partir de los datos del IVIE.

Gráfico 4.2

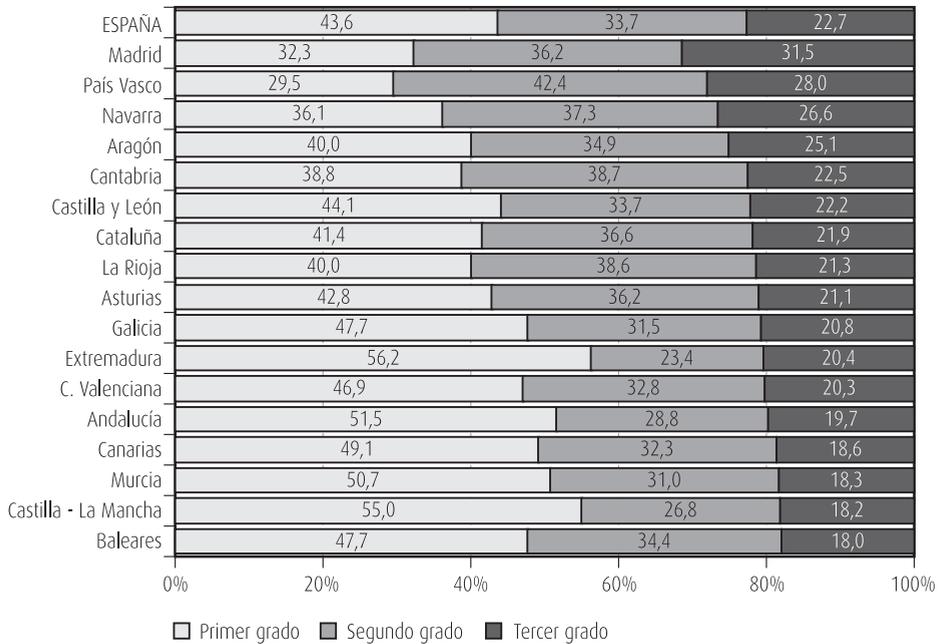
## DISTRIBUCIÓN GEOGRÁFICA DE LOS AÑOS MEDIOS DE EDUCACIÓN DE LA POBLACIÓN OCUPADA



Fuente: Cálculos propios a partir de los datos del IVIE.

Gráfico 4.3

## COMPOSICIÓN EDUCATIVA DE LOS OCUPADOS POR CCAA, 2006



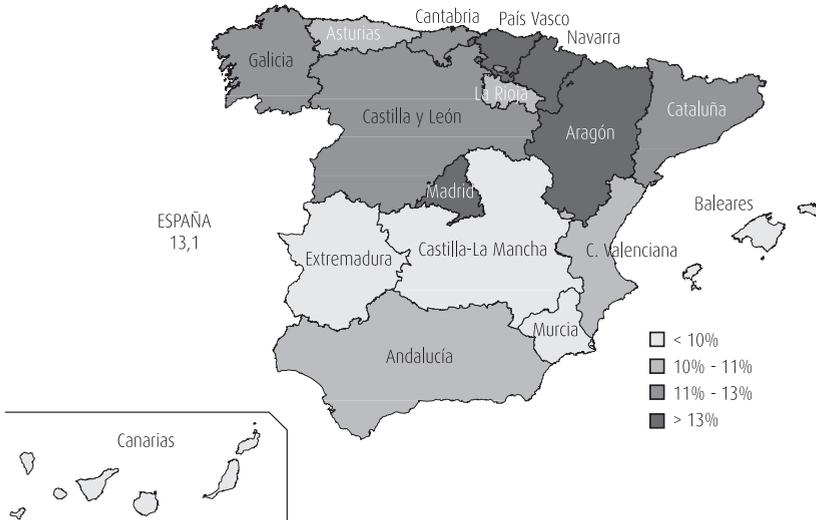
Fuente: Cálculos propios a partir de los datos del IVIE. Nota: Primer grado: Sin estudios, primaria y EGB; Segundo grado: FP1, FP2 y BUP; Tercer grado: antes de superior y superior.

El gráfico 4.3 muestra la proporción de ocupados por niveles educativos a nivel de comunidades autónomas. Se observa que en el 2006 alrededor de un 40% de los ocupados españoles presentan bajo nivel de cualificación y un 23% presentan alta cualificación. Por comunidad autónoma se tiene que en Madrid más de una tercera parte de su población ocupada presenta altos niveles de cualificación. En este orden de mayor cualificación se tiene también al País Vasco (28%) y Navarra (27%). Por su parte, entre las regiones con menor cualificación de sus ocupados se destacan Extremadura (56%) Castilla-La Mancha (55%), Andalucía (51%) y Murcia (50%). En cuanto a la distribución geográfica de los mayores niveles de capital humano de la población ocupada (gráfico 4.4) se observa que vuelve aparecer el nordeste del país con mayor cualificación de los ocupados

Ahora se pasa a hacer el análisis sectorial del capital humano. Teniendo en cuenta el comportamiento por comunidades autónomas, el gráfico 4.5 muestra los años medios de educación para los tres principales sectores económicos. Para el total de España se tiene que los ocupados en el sector industrial tienen en promedio 11,8 años de educación, los trabajadores de la construcción tienen una media de 10 años y los ocupados en el sector servicios presen-

Gráfico 4.4

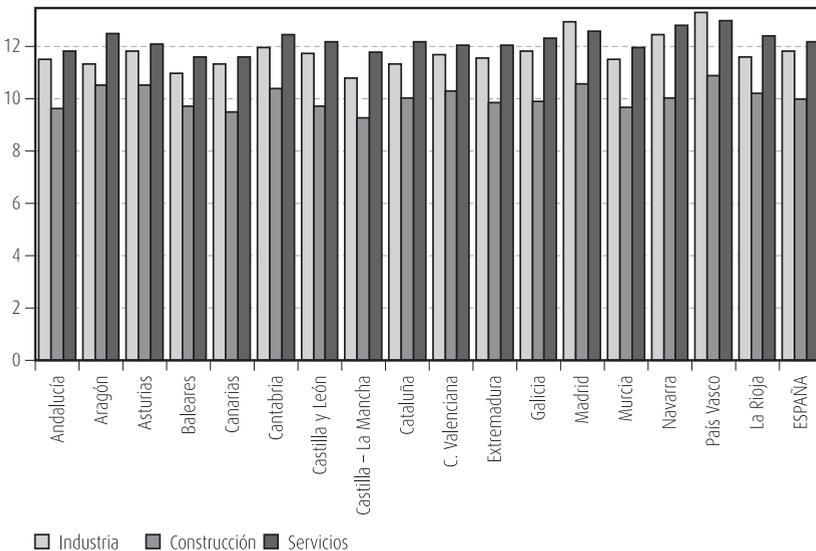
**PORCENTAJE DE LICENCIADOS POR CCAA, 2006**



Fuente: Cálculos propios a partir de los datos del IVIE.

Gráfico 4.5

**AÑOS MEDIOS DE EDUCACIÓN POR SECTORES Y CCAA, 2006**



Fuente: Cálculos propios a partir de los datos del IVIE.

tan en promedio 12.2 años de educación. A nivel de comunidades autónomas las diferencias son más marcadas. Por ejemplo, en el sector industrial las comunidades de País Vasco, Madrid y Navarra presentan los niveles medios de educación más altos: 13,3, 13, 9,2 y 12,5 años, respectivamente. Por su parte, en el sector de la construcción, las comunidades de País Vasco, Madrid, Aragón y Asturias presentan más de medio año adicional de educación por encima de la media nacional, a saber 10,9, 10,6, 10,6 y 10,5 años, respectivamente. En el sector de servicios las comunidades con mayores medias en educación son: País Vasco (13), Navarra (12,8), Madrid (12,6) y Aragón (12,5). Se nota la hegemonía del País Vasco y Madrid en la dotación de capital humano productivo y el carácter multisectorial que presentan.

#### ■ 4.5. EVOLUCIÓN DE LOS RENDIMIENTOS EDUCATIVOS EN EL PERÍODO 1995, 2002 Y 2006: DIFERENCIAS POR SEXO Y CC.AA.

Empleando la información de la Encuesta de Estructura Salarial (EES), se analizó la evolución temporal de los retornos educativos en España según sexo, nivel educativo y región para el periodo 1995-2006. Como se mencionó, esta encuesta ha presentado ajustes en su población objetivo a lo largo de su realización, por tanto para presentar resultados comparables en el tiempo, es preciso aplicar una serie de filtros y controles.

Los retornos fueron estimados para la población de ocupados por cuenta ajena que durante el mes de octubre se encontraban afiliados a la seguridad social, en centros de cotización de empresas de 10 y más empleados, de los sectores industria, construcción y servicios comprendidos en las secciones C a K de la CNAE-93. Adicionalmente, se seleccionó los individuos con edades entre 16 y 65 años. Las estimaciones son representativas para 17 comunidades autónomas.

Se aproximó los rendimientos educativos estimando ecuaciones de Mincer por dos vías, la primera, incluye los años de escolaridad como variable continua, mientras que la segunda, introduce dummies por nivel educativo.

Dado que la EES ofrece información sobre el máximo nivel educativo alcanzado por los individuos, la variable años de escolaridad fue imputada aproximando los siguientes valores por nivel:

- Hasta primaria completa (PRI): 4,5 años<sup>5</sup>.
- Secundaria obligatoria-ciclo 1 (EGB): 8 años.
- Formación profesional de grado medio (FP1): 9 años.
- Formación profesional de grado superior (FP2): 11 años.

<sup>5</sup> A pesar que la encuesta aporta información sobre los individuos sin estudios y con estudios primarios completos e incompletos, se decidió agrupar estas categorías adjudicando 4.5 años de escolaridad para el grupo, dado que la distribución de cada categoría era atípica para 2006 (véase Anexo 1).

Tabla 5.1

## ESTIMACIÓN DE LOS RETORNOS EDUCATIVOS PARA ESPAÑA

ECUACIONES CONTINUAS			
VARIABLE DEPENDIENTE: LOG SALARIO REAL POR HORA			
VARIABLES EXPLICATIVAS	1995	2002	2006
Escolaridad	0,0865 (222,7)	0,0791 (223,5)	0,0687 (178,1)
Experiencia	0,0503 (129,4)	0,0351 (94,7)	0,0295 (70,0)
Experiencia <sup>2</sup>	-0,00057 (-78,4)	-0,00035 (-49,0)	-0,00031 (-38,7)
Mujer	-0,2105 (-79,8)	-0,2593 (-115,0)	-0,2391 (-96,6)
Constante	0,7652 (119,6)	1,0724 (182,2)	1,2358 (181,6)
N.º Observaciones	139.208	149.232	111.859
R2	38,0%	36,3%	31,8%
ECUACIONES DISCRETAS			
VARIABLE DEPENDIENTE: LOG SALARIO REAL POR HORA			
VARIABLES EXPLICATIVAS	1995	2002	2006
EGB	0,1272 (42,8)	0,1003 (38,0)	0,0884 (30,1)
FP1	0,3468 (64,2)	0,3206 (74,2)	0,2889 (57,2)
FP2	0,4943 (111,4)	0,4540 (113,9)	0,3927 (85,9)
BUP	0,5529 (126,9)	0,4563 (107,6)	0,3973 (83,9)
DIP	0,8178 (131,0)	0,7681 (138,0)	0,6424 (104,9)
LIC	1,1347 (174,4)	1,0264 (186,5)	0,8953 (157,1)
Experiencia	0,0514 (132,5)	0,0385 (103,9)	0,0323 (76,3)
Experiencia <sup>2</sup>	-0,00062 (-85,9)	-0,00043 (-59,9)	-0,00037 (-45,9)
Mujer	-0,2049 (-78,5)	-0,2590 (-116,6)	-0,2407 (-98,4)
Constante	1,2427 (226,5)	1,4845 (310,6)	1,5867 (283,9)
N.º Observaciones	139.208	149.232	111.859
R2	39,9%	38,6%	33,8%

Estimación: MCO. Matriz robusta de varianzas y covarianzas. Entre ( ) estadísticos t. Fuente: Cálculos propios a partir de la EES.

- Secundaria postobligatoria-ciclo 2 (BUP): 12 años.
- Diplomados universitarios (DIP): 15 años.
- Licenciados (LIC): 17 años.

El cálculo de los rendimientos procede del ajuste de ecuaciones mincerianas del tipo:

$$\ln W = \beta_0 + \beta_1 \cdot egb + \beta_2 \cdot fp1 + \beta_3 \cdot fp2 + \beta_4 \cdot bup + \beta_5 \cdot dip + \beta_6 \cdot lic + \beta_7 \cdot bsexo + \beta_8 \cdot exp + \beta_9 \cdot exp^2 + \varepsilon$$

Como variable dependiente se tomó el logaritmo del salario mensual real por hora trabajada, y como variables explicativas se incluyen las *dummies* de nivel educativo (o los años

Tabla 5.2

### COMPORTAMIENTO DE LOS RENDIMIENTOS EDUCATIVOS PARA ESPAÑA: 1995-2006

NIVEL EDUCATIVO	1995	2002	2006	DIFERENCIA 1995-2006
EGB	4,2%	3,3%	2,9%	1,3
FP1	6,9%	6,4%	5,8%	1,2
FP2	8,2%	7,6%	6,5%	1,7
BUP	7,9%	6,5%	5,7%	2,2
DIP	8,2%	7,7%	6,4%	1,8
LIC	9,5%	8,6%	7,5%	2,0
<b>TOTAL</b>	<b>8,7%</b>	<b>7,9%</b>	<b>6,9%</b>	<b>1,8</b>

Fuente: Cálculos propios a partir de la EES.

de escolaridad), la *dummy* de sexo que toma valores de uno para mujeres, la experiencia y la experiencia al cuadrado, calculadas a partir de la experiencia potencial (edad-escolaridad-6).

El salario mensual se aproximó a partir de la agregación del salario base a octubre, los complementos salariales y las pagas extraordinarias. Se prefirió esta alternativa, frente a tomar el salario bruto anual, para evitar el posible ruido que pagas extemporáneas puedan suponer en la valoración de los salarios<sup>6</sup>. Respecto a las *dummies* de nivel educativo, primaria (PRI) se tomó como categoría de referencia (5 años), mientras que la categoría licenciados (LIC), agrupa todos los titulados superiores: licenciados y doctores.

El cálculo de los retornos a partir de la especificación continua de la ecuación de Mincer es directo para la muestra total o para subgrupos. Si se emplea la especificación por niveles educativos, el rendimiento se refiere siempre al rendimiento promedio con respecto a primaria, para la que se consideran 5 años de estudio. Para cada nivel se obtienen rendimientos con respecto a la categoría de referencia que es sin estudios. Por ejemplo, los rendimientos de la licenciatura se obtienen a través de:

$$\text{Rendimiento paso de estudios primarios a licenciado} = \frac{\beta_6}{17-5}$$

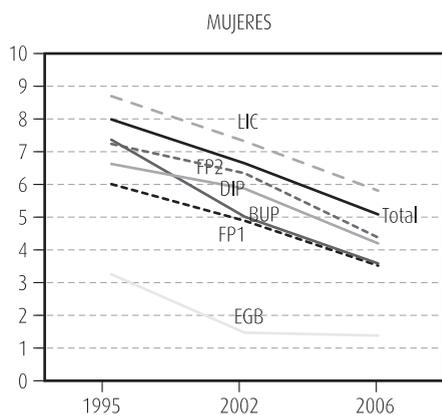
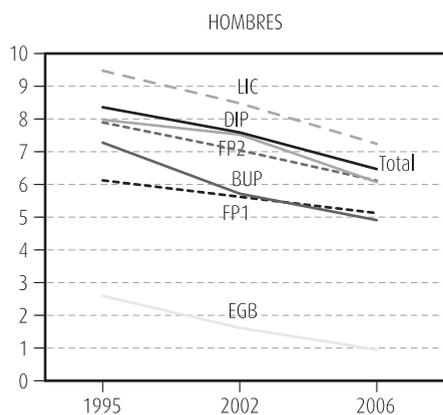
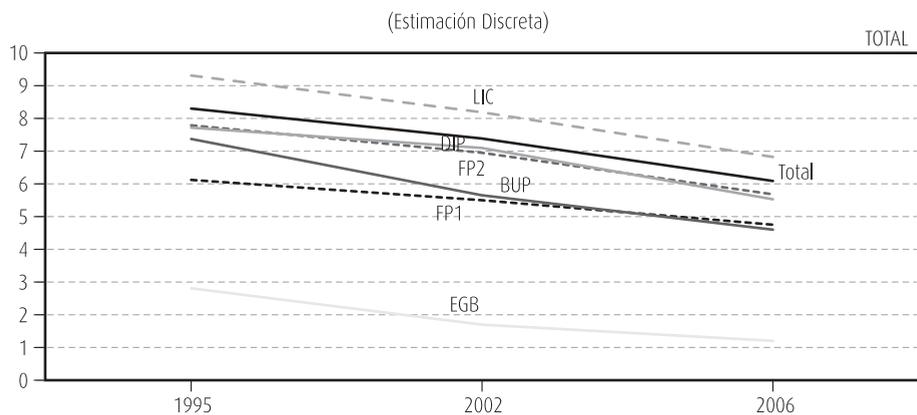
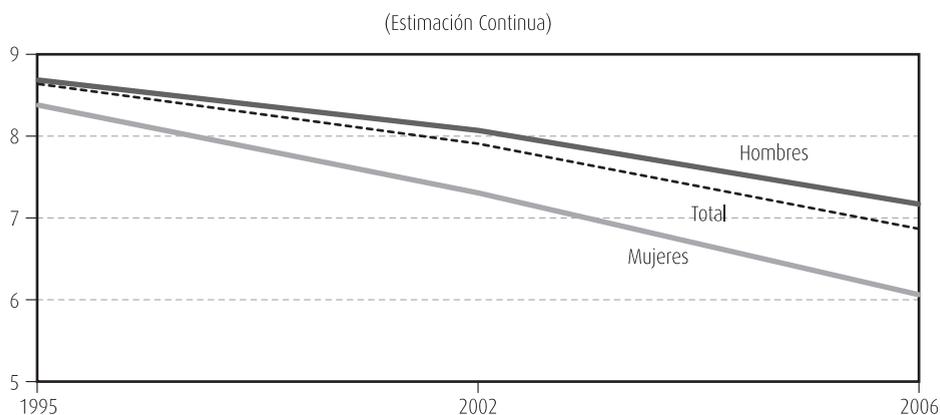
Los resultados de ambas estimaciones son presentadas en las tablas 5.1 y 5.2.

Como puede observarse, los rendimientos educativos han caído en un orden cercano a los 1,8 puntos porcentuales en la última década para el territorio español. Los niveles educativos más afectados por la reducción en los retornos son la secundaria posobligatoria y los estudios superiores (diplomados y licenciados).

<sup>6</sup> Se ajustó la estimación de los retornos para ambas variables, obteniendo un mejor ajuste para el salario mensual.

Gráfico 5.1

## EVOLUCIÓN DE LOS RENDIMIENTOS EDUCATIVOS SEGÚN SEXO



Fuente: Cálculos propios a partir de la EES.

#### ■ 4.5.1. Comportamiento de los rendimientos educativos según sexo

Al estimar estas funciones separando la muestra entre hombres y mujeres, se observa que el grupo más afectado por la reducción de los retornos educativos en la última década para España fueron las mujeres. Durante el año 1995 los hombres presentaban unos retornos a la educación de 8,7%, mientras que para las mujeres era de 8,4%. La brecha se amplió, teniendo una caída más pronunciada para las mujeres en el periodo 2006. Los retornos fueron de 7,2% y 6,1%, para hombres y mujeres respectivamente.

La estimación discreta, recoge un comportamiento similar. Los retornos para las mujeres, han sido menores para todos los niveles educativos y la caída es también más pronunciada durante el periodo analizado. El comportamiento gráfico de los retornos por sexo se presenta en el gráfico 5.1.

Los rendimientos de las mujeres son más bajos que los de los hombres. Evaluar el diferencial de rendimientos es difícil si bien una potencial explicación es que la oferta de educados ha crecido más rápido en el caso de las mujeres que en el de los hombres, lo que puede afectar negativamente al rendimiento. En cualquier caso, conviene resaltar que se habla de rendimientos (es decir, de diferencial salarial entre hombres o entre mujeres con distintos niveles educativos) y no de valores absolutos de salarios. Por tanto, el diferencial de rendimientos no es nunca interpretable en términos de discriminación.

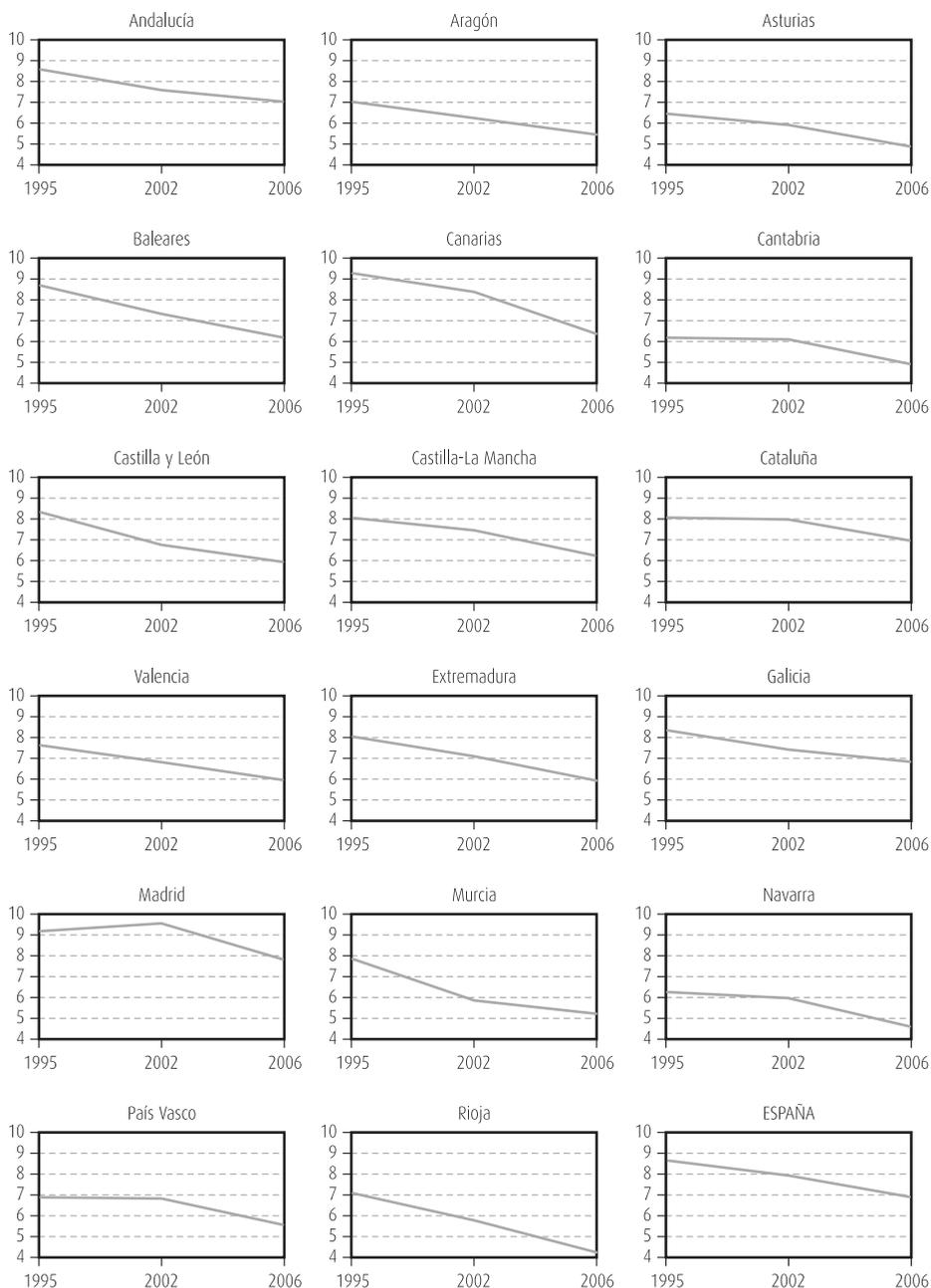
El diferencial de salarios sí podría ser interpretado en forma de discriminación. Pero para que esta interpretación fuese correcta las variables deberían estar medidas de forma adecuada. En particular, si bien la experiencia potencial puede ser una *proxy* razonable de la experiencia real en el caso de los hombres, en el caso de las mujeres, debido a los procesos de entrada y salida del mercado de trabajo, la experiencia potencial muy posiblemente sobrevalore la experiencia real, lo que justificaría un diferencial de salarios entre ambos colectivos a pesar de la inexistencia de discriminación.

La EES ofrece información sobre la antigüedad del individuo en el empleo actual, que es una medida directa de la experiencia, con la deficiencia de no incluir información sobre la experiencia previa que posee el individuo. Una forma obtener mayor evidencia respecto a la existencia de discriminación en el mercado laboral se podría aproximar estimando la ecuación de salarios incluyendo la antigüedad y la experiencia previa obtenida a partir de la diferencia entre experiencia potencial y la antigüedad<sup>7</sup>. El capítulo 4.3 presenta evidencia sobre el comportamiento de los salarios en la última década, y de los factores que lo explican en función de la evolución del capital humano en España.

<sup>7</sup> A pesar de que esta solución no está exenta del problema de usar la experiencia potencial, corrige de manera importante el posible sesgo presente en la información.

Gráfico 5.2

## EVOLUCIÓN DE LOS RENDIMIENTOS POR COMUNIDAD AUTÓNOMA



Fuente: Cálculos propios a partir de la EES.

#### ■ 4.5.2. Los retornos educativos por comunidad autónoma

Siguiendo el procedimiento anterior, se estimó los retornos para la muestra segmentada por CCAA, obteniendo los siguientes resultados (gráfico 5.2).

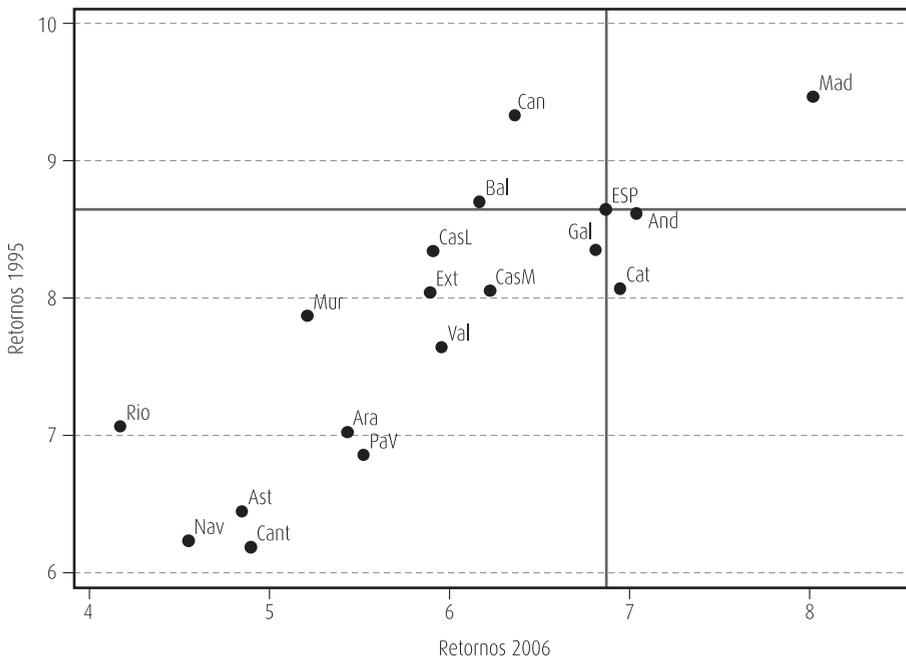
En general se identifica Madrid como la región con mayores retornos educativos, y regiones como Asturias, Cantabria y Navarra que presentan retornos muy bajos para los tres periodos analizados. En consonancia con los datos nacionales, los resultados por CCAA muestran también una reducción de rendimientos (gráfico 5.3).

Al comparar la evolución de los retornos a la educación entre 1995 y 2006, respecto a la media nacional, podemos observar que regiones como Cataluña y Andalucía ganaron posición relativa reubicándose en 2006 por encima de la media nacional, mientras que Canarias y Baleares perdieron posición en los últimos años. La reducción de los retornos para estas últimas es más importante, al igual que para Castilla y León, Murcia y Rioja.

A modo de conclusión, tenemos que el comportamiento de los retornos para España durante el periodo 1995-2006, ha presentado una indiscutible reducción que ha afectado

Gráfico 5.3

#### POSICIÓN RELATIVA DE LOS RETORNOS POR CCAA ENTRE 1995 Y 2006



Fuente: Cálculos propios a partir de la EES.

más a las mujeres, los individuos más educados y algunas regiones, entre las que están Baleares, Canarias, Castilla y León, Murcia y Rioja. Madrid se caracteriza por presentar unos retornos educativos muy por encima de la media nacional, resultado que seguramente está influenciado por su estructura productiva y la centralización de empleos con alta especialización en la capital española.

### ■ 4.5.3. Factores condicionantes de las diferencias salariales para España en la última década

En este apartado se pretende aportar evidencia empírica respecto al comportamiento de los salarios durante la última década y sus diferencias por sexo. Dado que la EES ofrece información de la antigüedad del individuo en el empleo actual, se cuantificará en que medida la reducción en la antigüedad media de la población ha afectado sobre una potencial caída en el salario medio de los asalariados en España.

En este sentido, se estima una ecuación minceriana de la forma:

$$\ln w_i = \beta_0 + \beta_1 \cdot edu_i + \beta_2 \exp v_i + \beta_3 \exp v_i^2 + \beta_4 ant_i + \beta_5 ant_i^2 + \beta_6 \cdot sex_i + \varepsilon_i$$

donde  $\ln w_i$  es el logaritmo del salario mensual real por hora trabajada,  $edu_i$  son los años de escolaridad,  $\exp v_i$  son los años de experiencia previa antes del empleo actual (experiencia potencial – antigüedad),  $ant_i$  son los años de antigüedad en el empleo actual y  $sex_i$  es la *dummy* de sexo que toma valores de uno si el individuo  $i$  es mujer. Los  $\beta$ 's son los coeficientes a estimar y  $\varepsilon_i$  es la perturbación aleatoria en el sentido convencional.

Los resultados de esta estimación se recogen en la tabla 5.3. Se estimó la función minceriana en la versión continua y discreta, para cada uno de los periodos analizados en el este estudio. En estas tablas, podemos confirmar una caída en el retorno a la educación y la experiencia entre 1995 y 2006. No obstante, a pesar que también se cae el retorno a la antigüedad su caída no es tan importante en magnitud. Adicionalmente, esta última cumple un papel muy importante en la consolidación de los retornos al capital humano. El coeficiente para la *dummy* que recoge el efecto de ser mujer, es negativo y significativo, lo que implica que no se rechaza la hipótesis que el salario de las mujeres es menor que el de los hombres, incluso considerando el mismo nivel educativo y antigüedad. Este efecto se mantiene en el tiempo. Finalmente la constante, que representa el salario que obtendría un individuo sin considerar el capital humano, ha aumentado en los últimos años, sugiriendo que es posible que el salario mínimo de participación haya aumentado. Las estimaciones para el caso continuo y el discreto aportan información similar.

A partir de la información contenida en la tabla 5.3, se elabora el gráfico 5.4. Este gráfico presenta el perfil salarial por nivel educativo según sexo para los periodos 1995 y 2006. La primera tabla contiene los perfiles para el caso convencional, cuando se estima a partir de la experiencia potencial, mientras que el segundo se elabora empleando las estimaciones de los retornos considerando la antigüedad en el empleo.

Tabla 5.3

### ESTIMACIÓN DE LA RENTABILIDAD DE LA EDUCACIÓN CONTROLANDO POR ANTIGÜEDAD\*

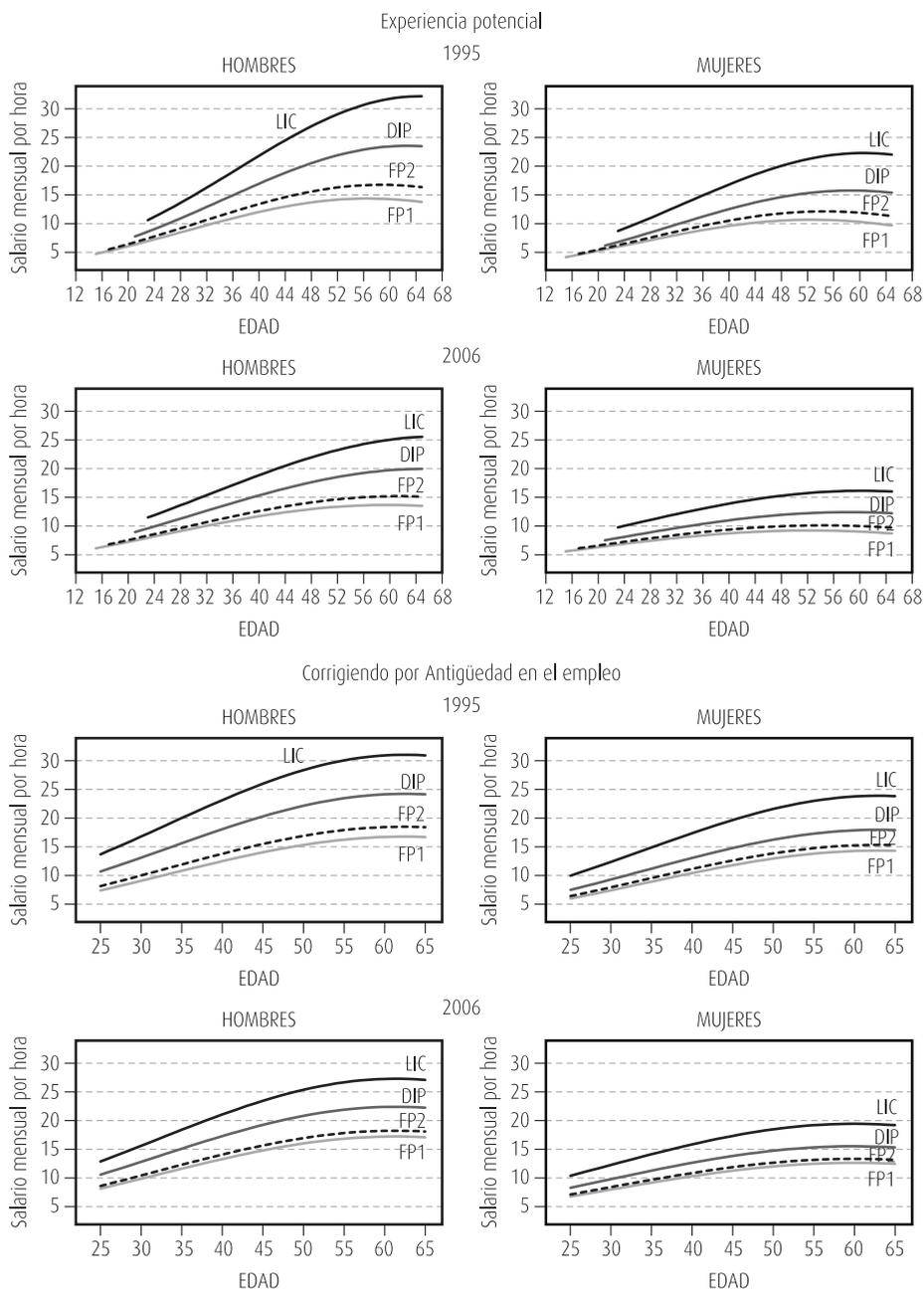
ECUACIÓN CONTINUA			
VARIABLE DEPENDIENTE: LOGARITMO DEL SALARIO REAL POR HORA			
VARIABLE	1995	2002	2006
Educación	0,0770 (218,9)	0,0656 (207,1)	0,0548 (160,4)
Experiencia previa	0,0199 (53,7)	0,0148 (42,2)	0,0096 (24,7)
Experiencia previa <sup>2</sup>	-0,00023 (-26,0)	-0,00017 (-21,9)	-0,00010 (-12,0)
Antigüedad	0,0465 (135,1)	0,0435 (131,9)	0,0401 (109,6)
Antigüedad <sup>2</sup>	-0,00062 (-55,4)	-0,00058 (-53,8)	-0,00056 (-48,4)
Mujer	-0,2185 (-84,1)	-0,2399 (-110,2)	-0,2216 (-92,8)
Constante	1,0801 (186,0)	1,3406 (246,8)	1,5056 (246,1)
N.º Observaciones	139.242	149.250	111.997
ECUACIÓN DISCRETA			
VARIABLE DEPENDIENTE: LOGARITMO SALARIO REAL POR HORA			
VARIABLE	1995	2002	2006
EGB	0,1130 (38,4)	0,0692 (26,2)	0,0548 (18,3)
FP1	0,3029 (56,5)	0,2372 (53,3)	0,2067 (40,5)
FP2	0,4219 (92,3)	0,3483 (86,6)	0,2815 (60,9)
BUP	0,4777 (118,3)	0,3458 (91,4)	0,2851 (68,1)
DIP	0,7315 (133,9)	0,6329 (129,7)	0,5064 (95,9)
LIC	1,0207 (191,1)	0,8699 (195,1)	0,7343 (155,4)
Experiencia previa	0,0191 (52,1)	0,0166 (48,0)	0,0118 (30,6)
Experiencia previa <sup>2</sup>	-0,00026 (-29,6)	-0,00024 (-30,5)	-0,00016 (-19,1)
Antigüedad	0,0445 (130,3)	0,0430 (132,8)	0,0397 (110,1)
Antigüedad <sup>2</sup>	-0,00060 (-54,5)	-0,00058 (-54,7)	-0,00055 (-47,8)
Mujer	-0,2134 (-83,2)	-0,2393 (-111,9)	-0,2228 (-94,7)
Constante	1,5312 (311,0)	1,7101 (384,9)	1,8036 (358,3)
N.º Observaciones	139.242	149.250	111.997

\*Estimación: MCO. Matriz robusta de varianzas y covarianzas. Entre ( ) estadísticos t. Fuente: Cálculos propios a partir de la EES.

Para la elaboración de los perfiles salariales corrigiendo por antigüedad, se supuso que el caso de un individuo típico que a los 25 años obtiene un empleo y permanece en él hasta su jubilación. La experiencia en este caso, viene definida por los años adquiridos según su nivel educativo hasta los 25 años. Se partió de estimaciones separadas para hombres y mujeres por año. Estas estimaciones nos permitirían comparar el comportamiento de los salarios para individuos con la misma dotación de capital humano que sólo difieren en el sexo y la temporalidad.

Gráfico 5.4

## DIFERENCIAS EN EL PERFIL SALARIAL POR SEXO Y NIVEL EDUCATIVO



Fuente: Cálculos propios a partir de la EES.

En el primer gráfico tomando experiencia potencial observamos diferencias importantes en el perfil salarial de hombres y mujeres por nivel educativo. Las mujeres obtienen en ambos periodos salarios inferiores a los hombres a pesar de su nivel educativo. Adicionalmente, se observa una caída generalizada en los salarios para ambos grupos y una reducción de la brecha de salarios para cada nivel educativo lo que confirma una reducción en retornos.

Diferencias en la antigüedad promedio de los individuos en el periodo analizado podrían contribuir a explicar diferencias en niveles salariales. No obstante, el gráfico que considera la antigüedad de manera homogénea para los distintos individuos confirma estas diferencias. Evidentemente esta corrección reduce la caída de los salarios durante el periodo analizado, y las diferencias por sexo, pero no las elimina. Las mujeres perciben un salario inferior al de los hombres a pesar de su nivel educativo, y si bien las diferencias son menores entre los periodos analizados, el efecto de reducción en los salarios persiste.

En cualquier caso, la pendiente, o sea, los retornos, influenciarán el comportamiento del perfil salarial para los individuos. La caída de los retornos afectará sin duda la evolución del salario de los individuos durante su ciclo vital, y la comparación de los salarios en ambos periodos dependerá del individuo al cual se haga referencia.

Para contrastar esta aparente reducción de los salarios e identificar que aspectos pueden haber influido sobre el comportamiento de los mismos durante el periodo analizado, se propone la aplicación del análisis *shift-share* para el comportamiento de los salarios en España. Esta es una herramienta de estática comparativa, que permite examinar las condiciones al principio y al final de dos periodos, descomponiendo los factores que condicionan su evolución.

Para ello se predice los salarios para los años 1995 y 2006, a partir de la siguiente expresión:

$$\begin{aligned} l\hat{w}_{95}(\hat{\beta}_{95}, \bar{X}_{95}) &= \hat{\beta}_{095} + \hat{\beta}_{195} \cdot \overline{edu}_{95} + \hat{\beta}_{295} \cdot \overline{expv}_{95} + \hat{\beta}_{395} \cdot \overline{expv^2}_{95} + \hat{\beta}_{495} \cdot \overline{ant}_{95} + \\ &\quad + \hat{\beta}_{595} \cdot \overline{ant^2}_{95} + \hat{\beta}_{695} \cdot \overline{sex}_{95} \\ l\hat{w}_{06}(\hat{\beta}_{06}, \bar{X}_{06}) &= \hat{\beta}_{006} + \hat{\beta}_{106} \cdot \overline{edu}_{06} + \hat{\beta}_{206} \cdot \overline{expv}_{06} + \hat{\beta}_{306} \cdot \overline{expv^2}_{06} + \hat{\beta}_{406} \cdot \overline{ant}_{06} + \\ &\quad + \hat{\beta}_{506} \cdot \overline{ant^2}_{06} + \hat{\beta}_{606} \cdot \overline{sex}_{06} \end{aligned}$$

que implica que el salario medio predicho en  $t$ , es función de los  $\beta$ 's estimados y de las características medias de la población para cada periodo. Así, si definimos:

$$\begin{aligned} \hat{\omega}_1 &\rightarrow l\hat{w}_{95} = f(\hat{\beta}_{95}, \bar{X}_{95}) \\ \hat{\omega}_2 &\rightarrow l\hat{w}_{06} = f(\hat{\beta}_{06}, \bar{X}_{06}) \\ \hat{\omega}_3 &\rightarrow l\hat{w}_{06} = f(\hat{\beta}_{95}, \bar{X}_{06}) \\ \hat{\omega}_4 &\rightarrow l\hat{w}_{06} = f(\hat{\beta}_{06}, \bar{X}_{95}) \end{aligned}$$

donde  $\hat{\omega}_1$  y  $\hat{\omega}_2$  representan el salario predicho en el sentido convencional para 1995 y 2006, respectivamente;  $\hat{\omega}_3$  y  $\hat{\omega}_4$  son los salarios predichos para 2006, suponiendo que los  $\beta$ 's no cambian o que las características no cambian respecto a 1995, respectivamente.

Podemos descomponer el cambio en el salario predicho entre 1995 y 2006, como el resultado del cambio en los coeficientes estimados, del cambio en las características de la población o del efecto mixto de estas relaciones, según:

$$\begin{aligned}\Delta TOTAL &= \hat{\omega}_2 - \hat{\omega}_1 = f(\hat{\beta}_{06}, \bar{X}_{06}) - f(\hat{\beta}_{06}, \bar{X}_{95}) \\ \Delta \bar{X} &= \hat{\omega}_3 - \hat{\omega}_1 = f(\hat{\beta}_{95}, \bar{X}_{06}) - f(\hat{\beta}_{95}, \bar{X}_{95}) \equiv \hat{\beta}_{95} \cdot (\bar{X}_{06} - \bar{X}_{95}) \\ \Delta \hat{\beta} &= \hat{\omega}_4 - \hat{\omega}_1 = f(\hat{\beta}_{06}, \bar{X}_{95}) - f(\hat{\beta}_{95}, \bar{X}_{95}) \equiv (\hat{\beta}_{06} - \hat{\beta}_{95}) \cdot (\bar{X}_{95}) \\ \Delta TOTAL &= \Delta X + \Delta \beta + \Delta Mixto\end{aligned}$$

donde,

$\Delta TOTAL$  representa el cambio en salario predicho entre ambos periodos,

$\Delta \bar{X}$  representa el cambio debido a variación en las características de la población

$\Delta \hat{\beta}$  representa el cambio que se debe a variaciones en los coeficientes estimados.

$\Delta Mixto$  representa el cambio por factores combinados de los anteriores.

Estos indicadores cabe expresarlos como:

$$\begin{aligned}\Delta \bar{X} &= \hat{\beta}_{195} \cdot (\overline{edu}_{06} - \overline{edu}_{95}) + \hat{\beta}_{295} \cdot (\overline{expv}_{06} - \overline{expv}_{95}) + \hat{\beta}_{395} \cdot (\overline{expv^2}_{06} - \overline{expv^2}_{95}) \\ &+ \hat{\beta}_{495} \cdot (\overline{ant}_{06} - \overline{ant}_{95}) + \hat{\beta}_{595} \cdot (\overline{ant^2}_{06} - \overline{ant^2}_{95}) + \hat{\beta}_{695} \cdot (\overline{sex}_{06} - \overline{sex}_{95}) \\ \Delta \bar{X} &= (\hat{\beta}_{006} - \hat{\beta}_{095}) + (\hat{\beta}_{106} - \hat{\beta}_{195}) \cdot \overline{edu}_{95} + (\hat{\beta}_{206} + \hat{\beta}_{295}) \cdot \overline{expv}_{95} + (\hat{\beta}_{306} - \hat{\beta}_{395}) \cdot \overline{expv^2}_{95} \\ &+ (\hat{\beta}_{406} - \hat{\beta}_{495}) \cdot \overline{ant}_{95} + (\hat{\beta}_{506} - \hat{\beta}_{595}) \cdot \overline{ant^2}_{95} + (\hat{\beta}_{606} - \hat{\beta}_{695}) \cdot \overline{sex}_{95}\end{aligned}$$

según lo cual, el cambio en el logaritmo del salario puede venir explicada por cambios en las características de la población según:

$\hat{\beta}_{195} \cdot (\overline{edu}_{06} - \overline{edu}_{95})$ : los años de escolaridad promedio

$\hat{\beta}_{295} \cdot (\overline{expv}_{06} - \overline{expv}_{95}) + \hat{\beta}_{395} \cdot (\overline{expv^2}_{06} - \overline{expv^2}_{95})$ : los años de experiencia previa promedio.

$\hat{\beta}_{495} \cdot (\overline{ant}_{06} - \overline{ant}_{95}) + \hat{\beta}_{595} \cdot (\overline{ant^2}_{06} - \overline{ant^2}_{95})$ : los años de antigüedad promedio.

$\hat{\beta}_{695} \cdot (\overline{sex}_{06} - \overline{sex}_{95})$ : el porcentaje de participación de las mujeres entre los asalariados.

o por puede estar influenciado por cambios en los coeficientes según:

$(\hat{\beta}_{106} - \hat{\beta}_{195}) \cdot \overline{edu}_{95}$ : Cambios en el retorno de la educación.

$(\hat{\beta}_{206} - \hat{\beta}_{295}) \cdot \overline{expv}_{95} + (\hat{\beta}_{306} - \hat{\beta}_{395}) \cdot \overline{expv^2}_{95}$ : cambios en el retorno a la experiencia previa.

$(\hat{\beta}_{406} - \hat{\beta}_{495}) \cdot \overline{ant}_{95} + (\hat{\beta}_{506} - \hat{\beta}_{595}) \cdot \overline{ant^2}_{95}$ : cambios en el retorno a la antigüedad.

$(\hat{\beta}_{606} - \hat{\beta}_{695}) \cdot \overline{sex}_{95}$ : cambio en la diferencia del ingreso de las mujeres respecto a los hombres.

$(\hat{\beta}_{006} - \hat{\beta}_{095})$ : cambio en el ingreso mínimo de los hombres.

Los resultados de estas diferencias estimadas se presentan en la tabla 5.4.

En la tabla 5.4, la predicción de los salarios nos muestra que el salario para el individuo promedio del año 1995 era mayor que para el individuo promedio del año 2006 (logaritmo del salario por hora pasa de 2,33 a 2,28, respectivamente)  $\omega_1 > \omega_2$ . No obstante, si la composición de los individuos no hubiese cambiado (permitiendo el cambio en los coeficientes) el salario aumentaría (logaritmo del salario por hora pasaría de 2,33 a 2,38)  $\omega_1 < \omega_4$ . Manteniendo los coeficientes estimados, y permitiendo cambios en las características de la población el salario se reduce (logaritmo del salario por hora pasaría de 2,33 a 2,23)  $\omega_1 > \omega_3$ .

Al descomponer estas diferencias en términos de cambios porcentuales según cada factor, tenemos que el salario cae en un 4.6%. No obstante, el cambio en las características reduce el salario predicho en un 9.5%, mientras que el cambio en los coeficientes

Tabla 5.4

### DIFERENCIAS SALARIALES EN ESPAÑA 1995-2006: ANÁLISIS SHIFT-SHARE

PREDICCIÓN DE LOS SALARIOS POR COMPONENTES				
VARIABLES/PREDICCIÓN SALARIAL	$\omega_1$	$\omega_2$	$\omega_3$	$\omega_4$
	$lW_{95}(\beta_{95}, X_{95})$	$lW_{06}(\beta_{06}, X_{06})$	$lW_{06}(\beta_{95}, X_{06})$	$lW_{06}(\beta_{06}, X_{95})$
Escolaridad	0,665	0,499	0,702	0,473
Experiencia previa	0,228	0,131	0,266	0,111
Antigüedad	0,409	0,224	0,262	0,350
Sexo (mujeres)	-0,054	-0,079	-0,078	-0,055
Ingreso mínimo hombres	1.080	1.506	1.080	1.506
<i>Logaritmo del Salario mensual por hora predicho</i>	2.327	2.281	2.232	2.385
DESCOMPOSICIÓN DE LOS FACTORES EXPLICATIVOS EN LAS DIFERENCIAS EN LOS LOGARITMOS DE LOS SALARIOS POR HORA				
VARIABLES/EFFECTO ESTIMADO	DIF. SALARIAL 1995-2006 ( $\Delta$ TOTAL)	CARACTERÍSTICAS ( $\Delta$ X)	COEFICIENTES ( $\Delta$ $\beta$ )	EFFECTOS CRUZADOS ( $\Delta$ MIXTO)
	$\omega_2 - \omega_1$	$\omega_3 - \omega_1$	$\omega_4 - \omega_1$	
Escolaridad	-16,5%	3,8%	-19,2%	-1,1%
Experiencia previa	-9,7%	3,8%	-11,6%	-1,9%
Antigüedad	-18,4%	-14,7%	-5,9%	2,2%
Sexo (% mujeres)	-2,5%	-2,4%	-0,1%	—
Ingreso mínimo hombres	42,5%	—	42,5%	—
<i>Logaritmo del Salario mensual por hora predicho</i>	-4,6%	-9,5%	5,7%	-0,8%

Fuente: Cálculos propios a partir de las EES.

aumenta el salario en un 5,7%. En este sentido, según las características de la población, la reducción en los años de antigüedad promedio reduce el salario predicho en un 14,7%, el aumento de la proporción de mujeres (que tienen en promedio un ingreso inferior), lo reduce en un 2,4%, mientras el aumento de los años de escolaridad y experiencia previa, lo aumenta en un 3,8%, respectivamente.

Respecto al cambio en los coeficientes, el factor con mayor efecto negativo son la caída del retorno a la educación y la experiencia, que reducen el salario en un 19.2% y 11.6%, respectivamente. La caída en el retorno a la antigüedad, lo reducen en un 5.9%. Sin embargo, un aumento en el salario básico (mínimo) para la población<sup>8</sup> en un 42.5% compensa el efecto de la caída que los retornos tienen sobre el salario. Este afecto combinado sería más acorde como una estandarización de los salarios, o sea, cada vez en el mercado laboral español cuenta menos la retribución a la educación en la fijación de los salarios. No obstante, el salario base en la última década efectivamente ha aumentado.

#### ■ 4.6. LA PROBABILIDAD DE EMPLEO Y LOS RENDIMIENTOS EDUCATIVOS

##### ■ 4.6.1. La consideración de las probabilidades de empleo en la evaluación de los rendimientos educativos

Los rendimientos educativos se ven afectados por la probabilidad de empleo. Con el fin de establecer la relación existente entre ambos para cada colectivo, procedemos a definir la probabilidad de empleo como:

$$p_i = \text{Probabilidad de empleo para el nivel educativo } i = 1 - \text{Probabilidad de paro para el nivel educativo } i$$

La probabilidad de paro para el nivel educativo  $i$  se obtiene a través de la estimación de un modelo *probit* que mide la probabilidad de paro de los individuos ubicados en este nivel a partir de la ecuación:

$$\text{Probabilidad de paro} = \Phi(\text{Edad}, \text{Edad}^2, \text{Sexo})$$

Se predice la probabilidad de paro del individuo desde que entra en el mercado laboral hasta que se retira. Si por ejemplo es licenciado y entra en el mercado laboral a los 23 años y se retira a los 65, se dispone de 43 predicciones de paro según la edad. Cabe obtener la probabilidad promedio de paro como una simple media aritmética de estas probabilidades anuales. La probabilidad de empleo se obtendría como:

$$p_i = 1 - \text{media aritmética de las probabilidades de paro para nivel educativo } i$$

<sup>8</sup> En este caso, se toma como referencia a los hombres, no obstante la diferencia entre hombres y mujeres no es muy importante: 0.1% menos para las mujeres.

Para calcular los rendimientos educativos se procede de la forma que seguidamente se detalla.

#### ■ 4.6.1.1. *Cálculo estándar sin probabilidad de empleo en el paso de estudios primarios a nivel de licenciado*

Siguiendo la especificación por niveles educativos presentada en la sección anterior, la ecuación estimada para obtener los retornos educativos es:

$$\ln W = \alpha + \beta_2 \cdot F_2 + \beta_3 \cdot F_3 + \beta_4 \cdot F_4 + \beta_5 \cdot F_5 + \beta_6 \cdot F_6 + \beta_7 \cdot F_7 + \gamma_1 \cdot Expe + \gamma_2 \cdot Expe^2 + \gamma_3 \cdot Sexo + \varepsilon$$

Sin introducir la corrección por no linealidades dado que se producen compensaciones, el valor esperado de los salarios de un licenciado es:

$$W_7 = \exp(\alpha + \beta_7 + \gamma_1 \cdot Expe + \gamma_2 \cdot Expe^2 + \gamma_3 \cdot Sexo)$$

El valor esperado de los salarios de un individuo con estudios primarios es:

$$W_1 = \exp(\alpha + \beta_1 + \gamma_1 \cdot Expe + \gamma_2 \cdot Expe^2 + \gamma_3 \cdot Sexo)$$

Si “ $n_7$ ” es el número de años requerido para licenciarse y “5” es el número de años requerido para completar estudios primarios, el rendimiento derivado del paso del nivel 1 al 7 viene dado por:

$$(1) \quad \rho_{1,7} = \frac{W_7 - W_1}{(n_7 - n_1) \cdot W_1} = \frac{\exp(\alpha + \beta_7 + \gamma_1 \cdot Expe + \gamma_2 \cdot Expe^2 + \gamma_3 \cdot Sexo) - \exp(\alpha + \beta_1 + \gamma_1 \cdot Expe + \gamma_2 \cdot Expe^2 + \gamma_3 \cdot Sexo)}{(n_7 - n_1) \cdot \exp(\alpha + \beta_1 + \gamma_1 \cdot Expe + \gamma_2 \cdot Expe^2 + \gamma_3 \cdot Sexo)}$$

Cuando la probabilidad de empleo no interviene, el rendimiento promedio puede aproximarse a través de:

$$(2) \quad \rho_{1,7} = \frac{W_7 - W_1}{(n_7 - n_1) \cdot W_1} \approx \frac{\beta_7}{(n_7 - n_1)}$$

No obstante, para que el cálculo resulte comparable cuando se hace intervenir la probabilidad de empleo, la expresión a utilizar debe ser la (1). Para computar la expresión (1) es posible predecir los salarios a lo largo del ciclo vital del individuo y referir el cálculo del salario al valor promedio del salario predicho. Cabe comparar los rendimientos educativos que se deducen de las expresiones (2) y (1) y comprobar que son relativamente similares.

Cabe observar que los rendimientos educativos pueden también expresarse a través de la prima salarial  $PW_7$ . Es decir, siendo  $PW_7 = \frac{W_7}{W_1}$ , se deduce:

$$(3) \quad \rho_{1,7} = \frac{W_7 - W_1}{(n_7 - n_1) \cdot W_1} = \frac{W_7 \cdot \left(1 - \frac{W_1}{W_7}\right)}{(n_7 - n_1) \cdot W_1} = \frac{W_7 \cdot \left(1 - \frac{1}{PW_7}\right)}{(n_7 - n_1) \cdot W_1} = \frac{1}{(n_7 - n_1)} \cdot PW_7 \cdot \left(1 - \frac{1}{PW_7}\right) =$$

$$= \frac{1}{(n_7 - n_1)} \cdot (PW_7 - 1)$$

Si la prima salarial es unitaria, el rendimiento educativo se hace cero. En realidad debería ser negativo, y el valor cero es consecuencia de que implícitamente se suponen vidas infinitas.

#### ■ 4.6.1.2. Cálculo con probabilidad de empleo en el paso de estudios primarios a nivel de licenciado

Si la probabilidad de empleo de un individuo con estudios primarios es " $\rho_1$ " y la probabilidad de empleo de un individuo con licenciatura es " $\rho_7$ ", para el cálculo de los rendimientos educativos se propone utilizar la siguiente expresión:

$$(3) \quad \rho_{1,7}^* = \frac{\rho_7 \cdot W_7 - \rho_1 \cdot W_1}{(n_7 - n_1) \cdot \rho_1 \cdot W_1} =$$

$$\frac{\rho_7 \cdot \exp(\alpha + \beta_7 + \gamma_1 \cdot Expe + \gamma_2 \cdot Expe^2 + \gamma_3 \cdot Sexo) - \rho_1 \cdot \exp(\alpha + \beta_1 + \gamma_1 \cdot Expe + \gamma_2 \cdot Expe^2 + \gamma_3 \cdot Sexo)}{(n_7 - n_1) \cdot \rho_1 \cdot \exp(\alpha + \beta_1 + \gamma_1 \cdot Expe + \gamma_2 \cdot Expe^2 + \gamma_3 \cdot Sexo)}$$

Obsérvese que la relación entre rendimientos educativos puede hallarse a través de la siguiente expresión:

$$\frac{\rho_{1,7}^*}{\rho_{1,7}} = \frac{\frac{W_7 \cdot \rho_7 - W_1 \cdot \rho_1}{W_1 \cdot \rho_1 \cdot (n_7 - n_1)}}{\frac{W_7 - W_1}{W_1 \cdot (n_7 - n_1)}} = \frac{\frac{\rho_7 \cdot \left(W_7 - W_1 \cdot \frac{\rho_1}{\rho_7}\right)}{\rho_1 \cdot W_1 \cdot (n_7 - n_1)}}{\frac{W_7 - W_1}{W_1 \cdot (n_7 - n_1)}} = \frac{\rho_7 \cdot \left(W_7 - W_1 \cdot \frac{\rho_1}{\rho_7}\right)}{\rho_1 \cdot W_1 \cdot (n_7 - n_1)} \cdot \frac{W_1 \cdot (n_7 - n_1)}{(W_7 - W_1)} =$$

$$= \frac{\rho_7 \cdot \left(W_7 - W_1 \cdot \frac{\rho_1}{\rho_7}\right)}{\rho_1 \cdot (W_7 - W_1)} = \frac{\rho_7 \cdot W_7 \cdot \left(1 - \frac{W_1}{W_7} \cdot \frac{\rho_1}{\rho_7}\right)}{\rho_1 \cdot W_1 \cdot \left(\frac{W_7}{W_1} - 1\right)} = \frac{\rho_7}{\rho_1} \cdot \frac{W_7}{W_1} \cdot \frac{\left(1 - \frac{W_1}{W_7} \cdot \frac{\rho_1}{\rho_7}\right)}{\left(\frac{W_7}{W_1} - 1\right)}$$

Denominando:

$$\text{Prima de probabilidad} = PP_7 = \frac{p_7}{p_1}$$

$$\text{Prima salarial} = PW_7 = \frac{W_7}{W_1}$$

Se deduce:

$$\frac{\rho_{1,7}^*}{\rho_{1,7}} = \frac{p_7}{p_1} \cdot \frac{W_7}{W_1} = \frac{\left(1 - \frac{W_1}{W_7} \cdot \frac{p_1}{p_7}\right)}{\left(\frac{W_7}{W_1} - 1\right)} = PP_7 \cdot PW_7 \cdot \frac{\left(1 - \frac{1}{PP_7 \cdot PW_7}\right)}{(PW_7 - 1)}$$

Es decir:

$$(4) \quad \rho_{1,7}^* = \rho_{1,7} \cdot PP_7 \cdot PW_7 \cdot \frac{\left(1 - \frac{1}{PP_7 \cdot PW_7}\right)}{(PW_7 - 1)} = \rho_{1,7} \cdot \frac{(PP_7 \cdot PW_7 - 1)}{(PW_7 - 1)}$$

Como cabía esperar, si la prima de probabilidad es unitaria, los dos rendimientos educativos se igualan. Es decir:  $\rho_{1,7}^* = \rho_{1,7}$

Según la expresión (4), si la prima de probabilidad aumenta, su efecto sobre el rendimiento educativo opera positivamente aumentando el numerador, mientras que el efecto derivado de la prima salarial es más elevado dado que aumenta el numerador y a la vez reduce el denominador.

En cualquier caso, el resumen de fórmulas a emplear para realizar este cálculo es el siguiente:

1) Ecuación salarial:

$$\ln W = \alpha + \beta_2 \cdot F_2 + \beta_3 \cdot F_3 + \beta_4 \cdot F_4 + \beta_5 \cdot F_5 + \beta_6 \cdot F_6 + \beta_7 \cdot F_7 + \gamma_1 \cdot Expe + \gamma_2 \cdot Expe^2 + \gamma_3 \cdot Sexo + \varepsilon$$

2) Modelo probit:

Probabilidad de paro a través estimación modelos probit =  $\Phi$  (Edad, Edad<sup>2</sup>, Sexo)

3) Probabilidad de empleo:

$p_i = 1 -$  Media aritmética de las probabilidades de paro para nivel educativo  $i$

4) Cálculo estándar de los rendimientos educativos:

$$\rho_{1,7} = \frac{W_7 - W_1}{(n_7 - n_1) \cdot W_1} \approx \frac{\beta_7}{(n_7 - n_1)}$$

5) Cálculo de los rendimientos educativos a través de los salarios promedio

Este cálculo debe ofrecer resultados muy similares a los precedentes:

$$\rho_{1,7} = \frac{W_7 - W_1}{(n_7 - n_1) \cdot W_1} =$$

$$= \frac{\exp(\alpha + \beta_7 + \gamma_1 \cdot Expe + \gamma_2 \cdot Expe^2 + \gamma_3 \cdot Sexo) - \exp(\alpha + \beta_1 + \gamma_1 \cdot Expe + \gamma_2 \cdot Expe^2 + \gamma_3 \cdot Sexo)}{(n_7 - n_1) \cdot \exp(\alpha + \beta_1 + \gamma_1 \cdot Expe + \gamma_2 \cdot Expe^2 + \gamma_3 \cdot Sexo)}$$

6) Cálculo de los rendimientos educativos incorporando la probabilidad de empleo:

$$\rho^*_{1,7} = \frac{p_7 \cdot W_7 - p_1 \cdot W_1}{(n_7 - n_1) \cdot p_1 \cdot W_1} =$$

$$= \frac{p_7 \cdot \exp(\alpha + \beta_7 + \gamma_1 \cdot Expe + \gamma_2 \cdot Expe^2 + \gamma_3 \cdot Sexo) - p_1 \cdot \exp(\alpha + \beta_1 + \gamma_1 \cdot Expe + \gamma_2 \cdot Expe^2 + \gamma_3 \cdot Sexo)}{(n_7 - n_1) \cdot p_1 \cdot \exp(\alpha + \beta_1 + \gamma_1 \cdot Expe + \gamma_2 \cdot Expe^2 + \gamma_3 \cdot Sexo)}$$

Considerar la probabilidad de empleo en el cálculo de los retornos educativos, implica evaluar la decisión a la que se enfrenta un individuo que decide educarse, si tuviera en cuenta el salario potencial que podría obtener si efectivamente logra emplearse o no. Gráficamente la relación sería la siguiente:

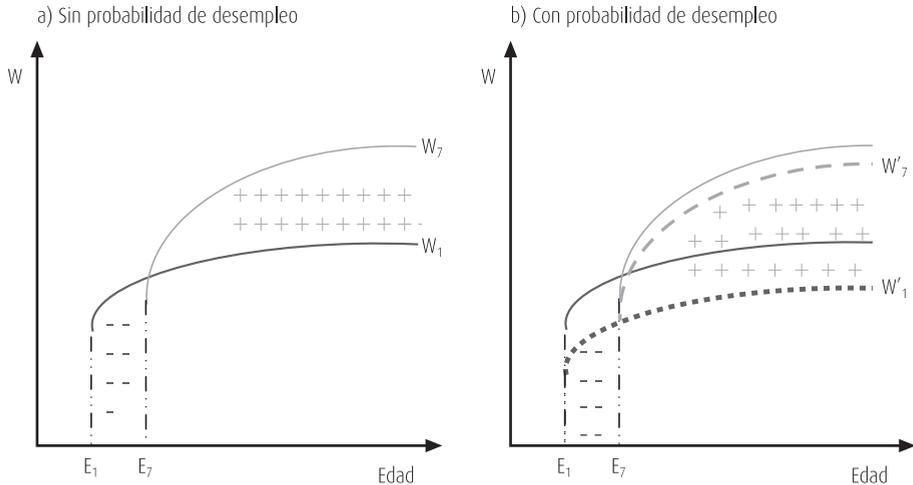
En el panel a) del gráfico 6.1, podemos observar la decisión a la que se enfrenta un individuo que decide educarse. La evolución del salario  $W$  a lo largo de su vida, vendrá condicionado por el nivel educativo que posea y la valoración que el mercado laboral hace de esta formación. Así, si un individuo que tiene nivel de estudios primarios decide hacer una licenciatura, el comportamiento de su salario a través de su ciclo vital, se desplazará de  $W_1$  (salario estudios primarios) a  $W_7$  (salario estudios licenciatura).

La decisión estará condicionada por la comparación de ambos flujos de ingresos, en el mismo sentido, como si de una inversión se tratara. La tasa de rentabilidad que ofrece obtener el título de licenciado, viene representada por la tasa de descuento que se obtiene al igualar el flujo positivo de ingresos adicionales que percibirá si decide educarse (signo +), con el flujo negativo que representa el costo de oportunidad de dejar de percibir el ingreso que le ofrece la educación primaria durante  $E_7 - E_1$ , o sea durante el tiempo que estaría estudiando y que retrasaría su entrada al mercado laboral (signo -).

Al considerar la probabilidad de desempleo (panel b del gráfico 6.1), los salarios esperados son ajustados según la probabilidad de emplearse (1-probabilidad desempleo). En el gráfico 6.1 se supone que la probabilidad de desempleo para los licenciados es menor que para los que tienen estudios primarios, lo que conlleva a que  $W_1$  se desplace en una mayor medida respecto de  $W_7$ . Al reestimar la rentabilidad de la educación para este nuevo escenario, se obtendrá una tasa de rentabilidad aún mayor para los licencia-

Gráfico 6.1

## VALORACIÓN DE LA RENTABILIDAD DE LA EDUCACIÓN CONSIDERANDO LA PROBABILIDAD DE DESEMPLEO



Fuente: Elaboración propia.

dos, dado el aumento de la corriente de positiva de ingresos potenciales en el flujo. No obstante, el retorno estimado considerando la probabilidad de desempleo, será mayor o menor dependiendo de cómo afecta el desempleo a los individuos según nivel educativo, en términos relativos.

### ■ 4.6.2. Efecto de la probabilidad de empleo sobre la estimación de los retornos en España

Para la estimación de los retornos educativos considerando la probabilidad de empleo, se utilizó la información de la Encuesta de Población Activa (EPA) y la Encuesta de Estructura Salarial (EES). A partir de la EPA se procedió al cálculo de la probabilidad de desempleo estimando modelos tipo probit de la forma:

$$\text{Probabilidad de paro} = \Phi(\text{Edad}, \text{Edad}^2, \text{Sexo}, \text{Región})$$

donde se define la probabilidad de paro, a partir del binario que toma valores de uno si está desempleado, o cero si está ocupado. En la muestra fueron seleccionados sólo los individuos de 16 a 65 años de edad, y la estimación se realizó segmentando la muestra para los niveles educativos:

- Primaria (PRI).
- Secundaria obligatoria (EGB).
- Formación profesional de grado medio (FP1).
- Formación profesional de grado superior (FP2).
- Secundaria postobligatoria (BUP).
- Diplomados universitarios (DIP).
- Titulados superiores (LIC).

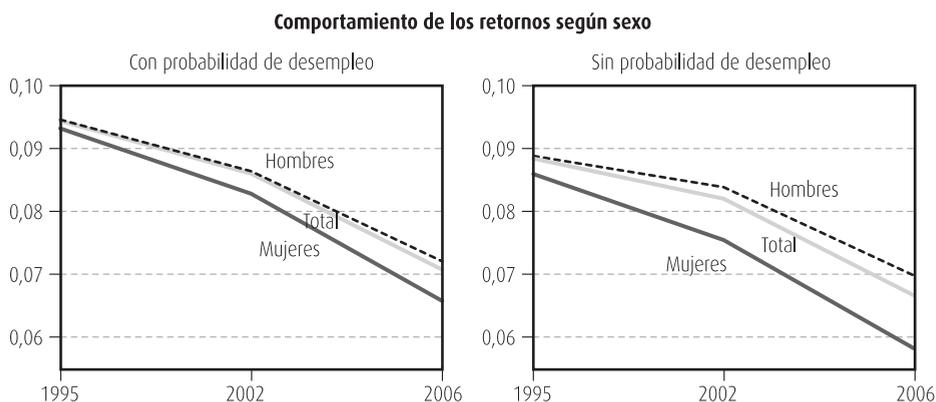
A partir de las probabilidades de paro predichas para cada individuo, se obtuvo la probabilidad de paro media para cada año según sexo, región y nivel educativo haciendo uso de los factores de expansión poblacional.

Se obtuvo las correspondientes probabilidades de empleo, y se imputó al salario de cada individuo típico en la EES, con el objeto de reestimar los retornos educativos. Los resultados de las ecuaciones estimadas se discuten a continuación.

Los gráficos 6.2 y 6.3 recogen el comportamiento de los retornos considerando la probabilidad de desempleo según sexo y nivel educativo, respectivamente. En primer lugar, los retornos para España estimados considerando la probabilidad de desempleo, son mayores que sin considerarla como se observa en el gráfico 6.2. Adicionalmente, se reduce la brecha existente entre los retornos de hombres y de mujeres, a pesar que las mujeres siguen

Gráfico 6.2

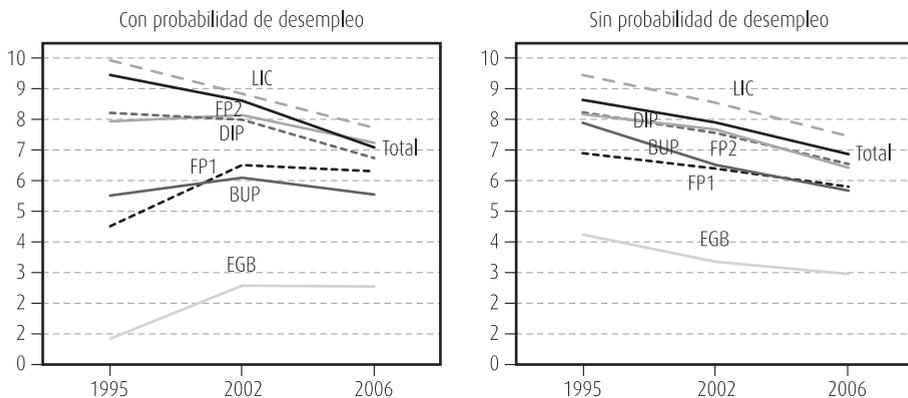
## EVOLUCIÓN DE LOS RETORNOS EDUCATIVOS SEGÚN SEXO CONSIDERANDO LA PROBABILIDAD DE DESEMPLEO



Fuente: Cálculos propios a partir de la EES y la EPA.

Gráfico 6.3

## EVOLUCIÓN DE LOS RETORNOS EDUCATIVOS SEGÚN NIVEL EDUCATIVO CONSIDERANDO LA PROBABILIDAD DE DESEMPLEO



Fuente: Cálculos propios a partir de la EES y la EPA.

presentando un retorno a la educación inferior al de los hombres, y que su caída en los últimos años es también mayor.

Respecto al retorno estimado por nivel educativo (gráfico 6.3), cabe resaltar que para el año 2006 los retornos con o sin probabilidad de desempleo difieren sólo ligeramente, pero no es así para 1995, cuando la economía presentaba altos niveles de desempleo. Como se explicó en la introducción de este capítulo, en presencia de altos niveles de desempleo, la rentabilidad de la educación dependerá de la probabilidad de estar ocupado.

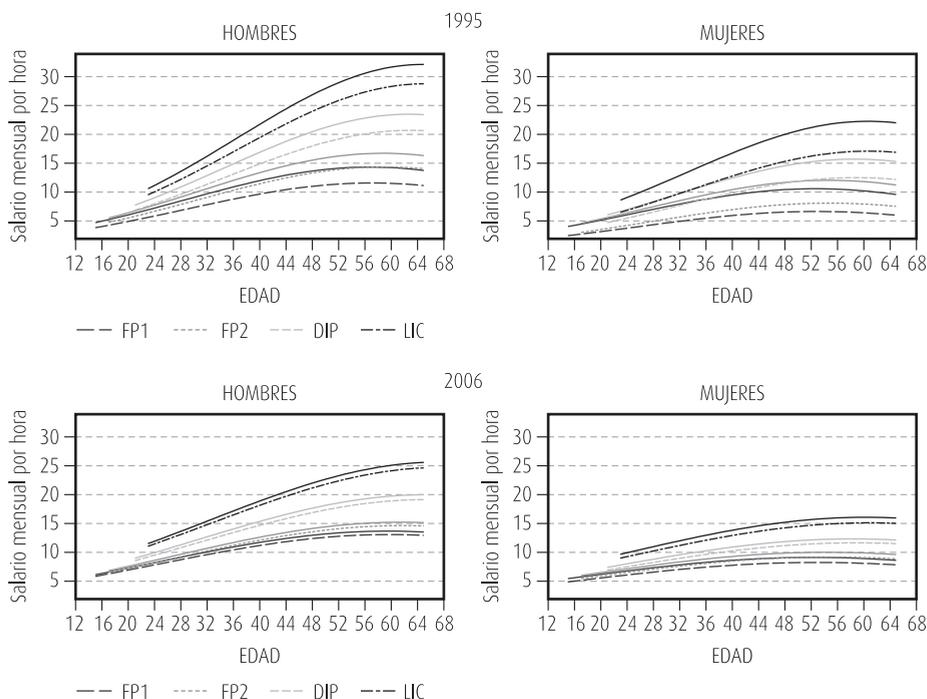
En periodos de recesión, la educación actúa como un mecanismo natural de selección usado por los empleadores frente a la larga cola de espera. Son los individuos más educados quienes tienen mayor probabilidad de emplearse. Este comportamiento se observa claramente en el gráfico 6.3, que presenta un aumento en los retornos de los licenciados, al pasar de la estimación con versus sin probabilidad de desempleo para 1995, frente al resto de niveles educativos.

El gráfico 6.4 presenta el perfil salarial según sexo para los individuos más educados, considerando su evolución con y sin probabilidad de desempleo. Este gráfico nos presenta varios aspectos de interés:

- Se confirma que el salario por hora percibido por las mujeres es menor al de los hombres en ambos periodos, comparando para distintos niveles educativos.
- Atendiendo a los perfiles salariales, se observa una reducción en los salarios para España entre 1995 y 2006, para ambos sexos.

Gráfico 6.4

## PERFIL SALARIAL CON Y SIN PROBABILIDAD DE DESEMPLEO



\*Las líneas punteadas corresponden al perfil salarial predicho según la probabilidad de desempleo. Fuente: Cálculos propios a partir de la EES y la EPA.

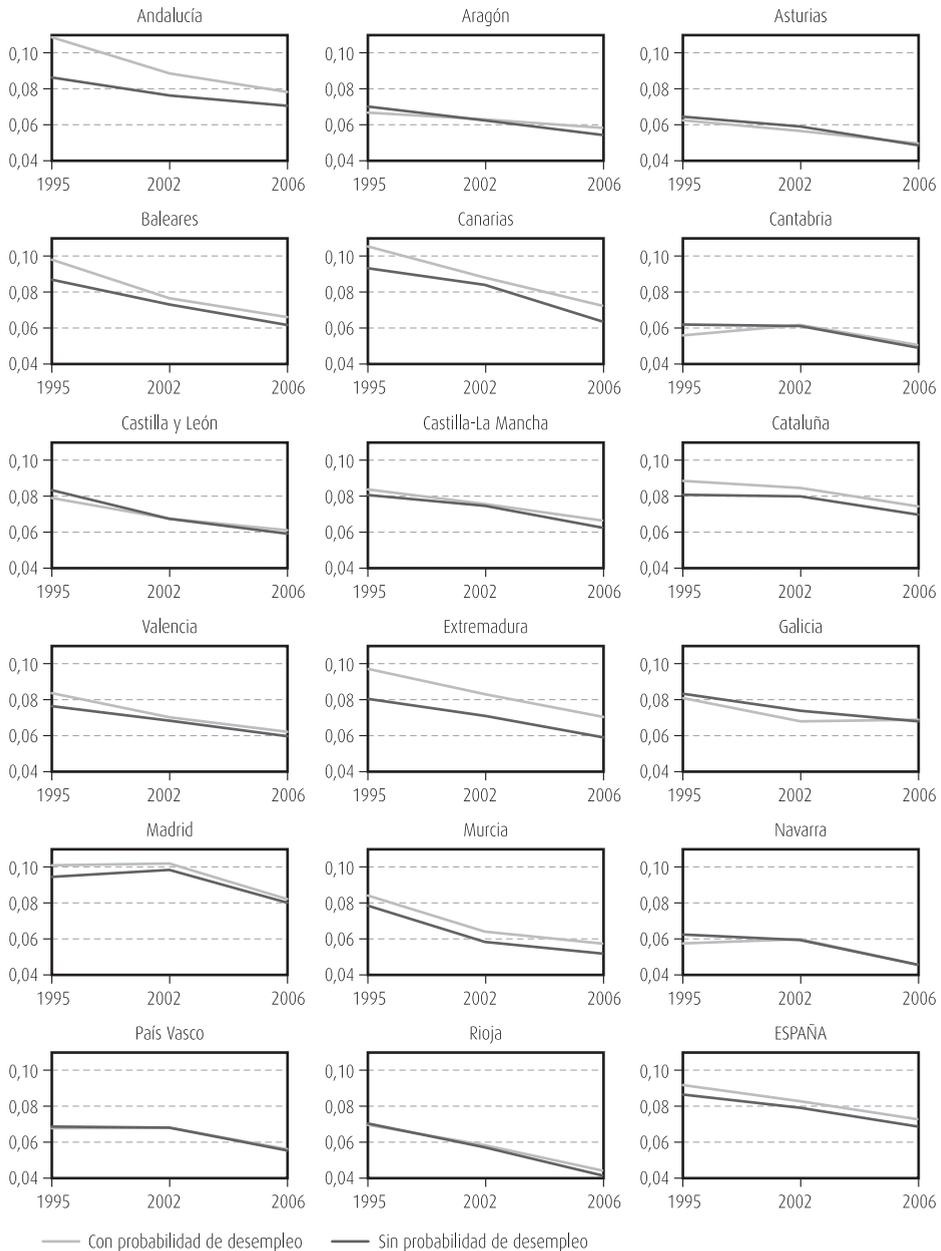
- La brecha salarial entre los niveles educativos medios, cada vez se reduce más, siendo en algunos casos nula o imperceptible.
- La prima que representa realizar estudios superiores es también menor.
- Las diferencias entre los salarios afectados por la probabilidad de desempleo (líneas discontinuas), frente a los salarios sin considerar la probabilidad de desempleo (líneas continuas), son mayores para 1995 frente a 2006.

El papel que desempeña la educación en el aparato productivo, es fundamental para liderar las decisiones de la oferta y la demanda de trabajo respecto a la dotación de capital humano. Obsérvese por ejemplo, el comportamiento de los retornos sin probabilidad de desempleo que presenta la secundaria posobligatoria (BUP) en el gráfico 6.3. Su caída es bastante significativa reposicionándose respecto a los ciclos de formación profesional. Esto sugiere que el BUP más que un tipo de formación relevante para el mercado de trabajo, se ha convertido en una puerta de acceso a los estudios superiores en los últimos años.

Gráfico 6.5

## COMPORTAMIENTO DE LOS RETORNOS SEGÚN COMUNIDAD AUTÓNOMA CONSIDERANDO LA PROBABILIDAD DE DESEMPLEO

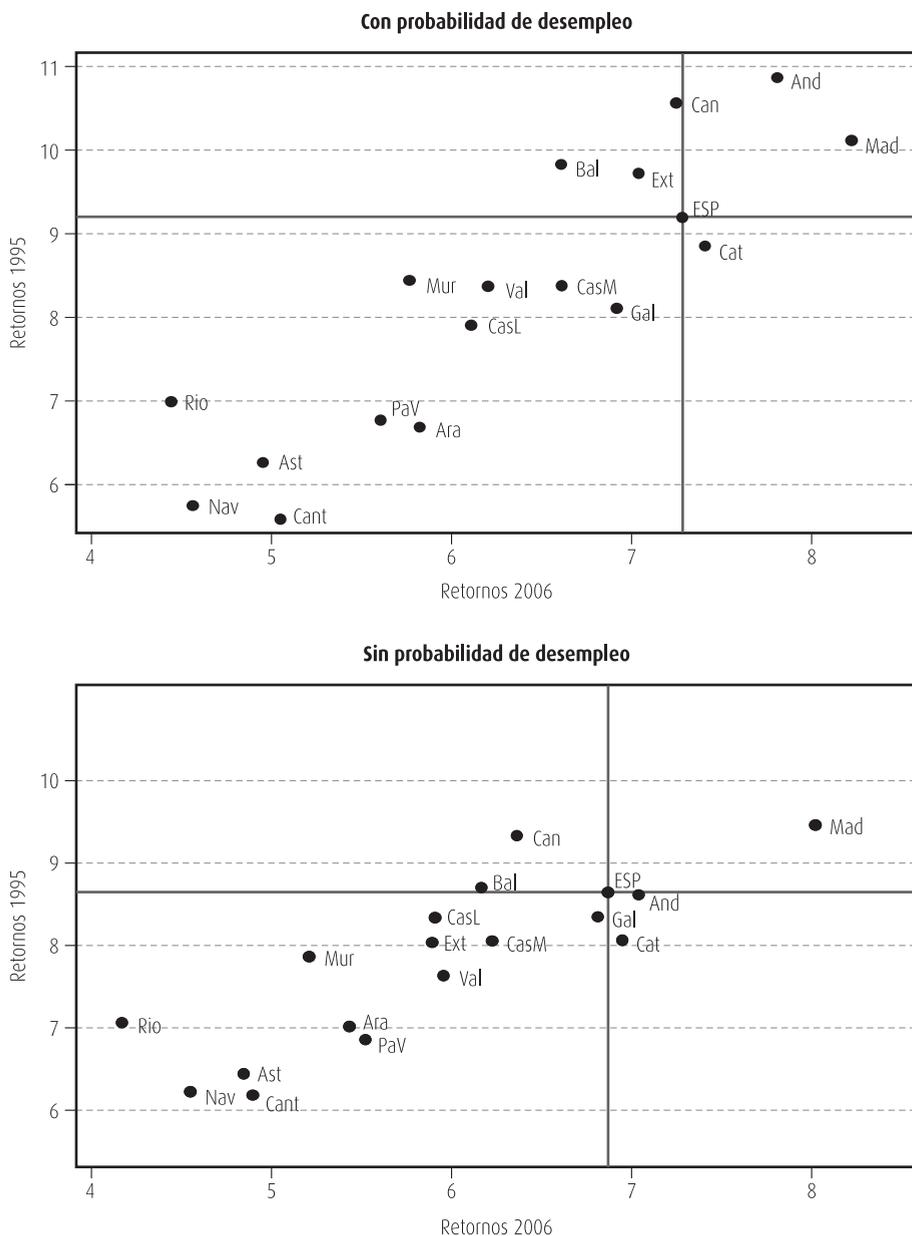
### Comportamiento de los retornos por comunidad autónoma



Fuente: Cálculos propios a partir de la EES y la EPA.

Gráfico 6.6

## POSICIÓN RELATIVA DE LOS RETORNOS EN LAS REGIONES CONSIDERANDO LA PROBABILIDAD DE DESEMPLEO



Fuente: Cálculos propios a partir de la EES y la EPA.

El gráfico 6.5 recoge las diferencias en la evolución de los retornos según región. Como cabría esperarse, son aquellas regiones con mayores tasas de desempleo las que mayores diferencias presentan. El gráfico 6.6 muestra la posición relativa de los retornos entre 1995 y 2006, estimados con y sin probabilidad de desempleo. Este gráfico confirma que los retornos a la educación cambian en mayor medida para Andalucía, Canarias y Extremadura en 1995 al considerar la probabilidad de desempleo.

A modo de conclusión, al considerar la probabilidad de desempleo para el cálculo de los retornos educativos en España, tenemos que los cambios en la estimación son más evidentes para 1995, periodo con altos niveles de desempleo, frente a 2006, cuando la economía se encontraba claramente en fase expansiva. Para este periodo respecto a la estimación sin considerar la probabilidad de desempleo, se puede resaltar la reducción de la brecha en los retornos por sexos, la reducción significativa en los retornos para los niveles educativos inferiores, y un mayor retorno para regiones como Andalucía, Canarias o Extremadura.

Respecto al comportamiento por regiones, es posible que el desempleo se concentre sobre la población con estudios inferiores, generando una brecha importante para los niveles educativos superiores, lo cual hace más rentable la educación. En estas regiones con menor proporción de individuos con estudios superiores, la oferta relativa de educados puede hacer rentable la decisión de educarse.

#### ■ 4.7. EL SESGO DE HABILIDAD Y LA FORMACIÓN DE UN PSEUDO PANEL

El análisis de la relación entre la escolaridad y los ingresos como forma de determinar el retorno a la inversión en educación (Mincer, 1974), puede presentar sesgos de habilidad, endogeneidad y errores de medición de la variable “escolaridad”. Los dos primeros tienden a sobreestimar los retornos, mientras que el tercero tiende a atenuarlos (Blackburn y Neumark, 1991, 1995; Behrman y Rosenzweig, 1999). El sesgo de habilidad y endogeneidad se encuentran estrechamente relacionados y en ocasiones pueden entenderse como el mismo problema. Puede darse el caso en el que el término de perturbación en la ecuación de ingresos refleje, entre otros factores, la habilidad innata de los individuos, lo que traería consigo la aparición del denominado sesgo de habilidad en la medida en que los individuos más hábiles sean aquellos que obtengan mayores niveles de escolaridad. En este caso la perturbación aleatoria y la variable de escolaridad estarían correlacionadas y, en consecuencia, aparecería el sesgo de endogeneidad. A pesar de que la habilidad innata no esté directamente correlacionada con la escolaridad, el sesgo de endogeneidad se plantea si el salario y la decisión de educarse se determinan conjuntamente. Con objeto de clarificar ideas, denominando  $W$  al salario,  $S$  a los años de escolaridad y  $Z$  y  $X$  al resto de variables exógenas, tenemos un sistema de dos ecuaciones:

$$\begin{aligned} W &= \beta \cdot S + \gamma' \cdot Z + u \\ S &= \phi \cdot W + \delta' \cdot X + v \end{aligned}$$

Resolviendo el sistema de ecuaciones y obteniendo la forma reducida es inmediato comprobar que  $u$  de la ecuación de salarios está influido por las perturbaciones aleatorias  $u$  y  $v$ , y por tanto que  $S$  de la ecuación de salarios está correlacionado con  $u$ , lo que provocará que la estimación por MCO sea inadecuada. No obstante, en la medida en que se disponga de instrumentos válidos que muestren una elevada correlación  $S$  con  $y$  que sean independiente de la perturbación aleatoria  $u$  (en definitiva, las variables exógenas  $X$  y  $Z$ ), la aplicación de Variables Instrumentales solucionará a la vez los sesgos de habilidad y de endogeneidad.

Un recurso para tratar el posible sesgo de habilidad en las ecuaciones de ingreso es intentar incorporar factores asociados con las características intrínsecas de los individuos o de su entorno familiar. Las características intrínsecas son aquellos factores inobservables que están relacionados con las habilidades innatas. Por su parte, el entorno familiar tienen que ver con el *background* familiar, es decir con las características de los padres, como su nivel educativo o tipo de trabajo, o la composición familiar. La incorporación de los anteriores factores depende de la disponibilidad de información. Una forma de tener en cuenta los efectos de la habilidad innata de los individuos en la ecuación de salarios, es disponer de información de individuos en el tiempo (datos panel) y aplicar el enfoque de efectos fijos (los cuales caracterizarían los elementos inobservables de cada individuo). Sin embargo, la información de tipo panel para individuos es poco común, a la vez que el anterior enfoque es poco utilizado entre otras razones porque la dimensión temporal del panel suele ser reducida y no permite observar variabilidad en el nivel educativo de los individuos encuestados. En tal caso, el nivel educativo de los individuos queda subsumido en el efecto fijo. Por su parte, las variables de *background* familiar son comúnmente más utilizadas para tratar el sesgo de habilidad a partir de la técnica de variables instrumentales y es más común que aparezcan en las encuestas (Blackburn y Neumark, 1991 y 1995; Uusitalo, 1999; Brunello y Miniaci, 1999; Butcher y Case, 1994).

En nuestro caso, como ya se comentó, la EES es una encuesta de corte transversal y no incluye información sobre el *background* familiar de los individuos. Ante esta limitación se va a seguir la metodología de pseudo-panel o panel sintéticos para tratar el sesgo de habilidad en las ecuaciones de ingreso. La metodología de pseudo-panel, tal como fue propuesto por Deaton (1985) y Browning, Deaton e Irish (1985), y luego discutido por Moffitt (1993), supera la dificultad de estudiar las dinámicas del ciclo de vida causada por la no disponibilidad de información de individuos en diferentes periodos de tiempo (Attanazio, 1997).

El procedimiento de pseudo-panel implica dividir a los individuos de la muestra en grupos definidos por una característica común (o cohorte), sea educación o año de nacimiento, y se siguen estos grupos a lo largo de las encuestas que se tengan. En este trabajo se han definido 12 cohortes teniendo en cuenta el año de nacimiento, el intervalo de las cohortes es de 3 años y estas cohortes se han seguido a lo largo de las EES de 1995, 2002 y 2006. La selección del intervalo que define a la cohorte es arbitraria. Un intervalo corto podría reducir la heterogeneidad entre los grupos además de incrementar el número de observaciones que se tengan. Lo anterior para nuestro caso no ha implicado una reducción significativa en el

Tabla 7.1

## DEFINICIÓN DE LAS COHORTES

COHORTE	FECHA DE NACIMIENTO		FECHA DE LA ENCUESTA		
	INTERVALO	PUNTO MEDIO	1995	2002	2006
			EDAD		
Cohorte 1	1977-1979	1978	16-18	23-25	27-29
Cohorte 2	1974-1976	1975	19-21	26-28	30-32
Cohorte 3	1971-1973	1972	22-24	29-31	33-35
Cohorte 4	1968-1970	1969	25-27	32-34	36-38
Cohorte 5	1965-1967	1966	28-30	35-37	39-41
Cohorte 6	1962-1964	1963	31-33	38-40	42-44
Cohorte 7	1959-1961	1960	34-36	41-43	45-47
Cohorte 8	1956-1958	1957	37-39	44-46	48-50
Cohorte 9	1953-1955	1954	40-42	47-49	51-53
Cohorte 10	1950-1952	1951	43-45	50-52	54-56
Cohorte 11	1947-1949	1948	46-48	53-55	57-59
Cohorte 12	1944-1946	1945	49-51	56-58	60-62

número de individuos por cada cohorte, en comparación a un intervalo de 5 años. En las tablas 7.1 y 7.2 se muestran las definiciones de las cohortes y el número de observaciones para cada cohorte. El análisis se realiza separadamente para hombres y mujeres.

De la tabla 7.2 se observa que por cohorte y años se tiene un número significativo de observaciones, más de 800 observaciones en promedio por cada cohorte, lo cual es adecuado para evitar el problema de errores de medida en las variables que se genera por tener diferentes individuos en cada sección cruzada (heterogeneidad y poca representatividad de las variables) (Browning, Deaton, Irish, 1985; Antman y McKenzie, 2005). Por otro lado, en la medida en que cada observación individual es sustituida por una media de individuos seleccionados de forma aleatoria utilizando como criterio el año de nacimiento, cabe esperar que el sesgo de habilidad desaparezca. En efecto, con objeto de centrar ideas, siendo el salario, el conjunto de variables explicativas y la perturbación aleatoria, se tiene:

$$\ln(W_{ij}) = \beta' \cdot X_{ij} + u_{ij}$$

$$u_{ij} = h_i + \varepsilon_{ij}$$

en donde "i" hace referencia al individuo y "j" a la cohorte. En este caso la perturbación aleatoria  $u_{ij}$  es la suma de la habilidad innata inobservable de cada individuo  $h_i$  y un componente puramente aleatorio  $\varepsilon_{ij}$ .

Al formar un pseudo panel cada individuo es sustituido por la media de individuos de la cohorte de forma que se tiene:

Tabla 7.2

## NÚMERO DE OBSERVACIONES POR COHORTE

COHORTE	MUESTRA DE HOMBRES			TOTAL	PROMEDIO
	AÑO				
	1995	2002	2006		
1978	582	7,605	6,668	14,855	4,951.7
1975	2,299	9,931	7,538	19,768	6,589.3
1972	5,395	10,083	7,005	22,483	7,494.3
1969	8,261	9,766	6,671	24,698	8,232.7
1966	9,858	9,372	6,085	25,315	8,438.3
1963	10,134	8,557	5,825	24,516	8,172
1960	9,644	7,990	5,196	22,830	7,610
1957	9,266	7,343	4,741	21,350	7,116.7
1954	8,641	6,573	3,920	19,134	6,378
1951	8,445	5,995	3,448	17,888	5,962.7
1948	8,772	5,443	2,860	17,075	5,691.7
1945	7,590	4,126	1,949	13,665	4,555
<b>Total</b>	<b>88,887</b>	<b>92,784</b>	<b>61,906</b>	<b>243,577</b>	<b>811,92.3</b>
COHORTE	MUESTRA DE MUJERES			TOTAL	PROMEDIO
	AÑO				
	1995	2002	2006		
1978	251	4,868	4,613	9,732	3,244
1975	1,483	6,240	4,821	12,544	4,181.3
1972	3,289	5,918	4,239	13,446	4,482
1969	4,368	4,820	3,597	12,785	4,261.7
1966	4,243	4,086	3,155	11,484	3,828
1963	3,474	3,638	2,856	9,968	3,322.7
1960	3,410	3,286	2,563	9,259	3,086.3
1957	3,132	3,079	2,242	8,453	2,817.7
1954	2,583	2,453	1,650	6,686	2,228.7
1951	1,988	1,819	1,175	4,982	1,660.7
1948	1,565	1,361	835	3,761	1,253.7
1945	1,195	905	532	2,632	877.3
<b>Total</b>	<b>30,981</b>	<b>42,473</b>	<b>32,278</b>	<b>105,732</b>	<b>35,244</b>

Fuente: Cálculos propios a partir de las EES.

$$\ln(W_j) = \frac{1}{N_j} \sum_{i=1}^{N_j} \ln(W_{ij})$$

$$X_j = \frac{1}{N_j} \sum_{i=1}^{N_j} X_{ij}$$

$$U_j = \frac{1}{N_j} \sum_{i=1}^{N_j} U_{ij} = \frac{1}{N_j} \sum_{i=1}^{N_j} (h_i + \varepsilon_{ij}) = \frac{1}{N_j} \sum_{i=1}^{N_j} (\varepsilon_{ij}) + \frac{1}{N_j} \sum_{i=1}^{N_j} (h_i) \approx \frac{1}{N_j} \sum_{i=1}^{N_j} (\varepsilon_{ij}) + \mu = \varepsilon_j + \mu$$

en donde  $N_j$  es el número de individuos de la cohorte  $j$ . Si la habilidad innata se distribuye de forma homogénea entre los individuos de cada cohorte, la media de habilidades innatas de cada cohorte tenderá a una media común  $\mu$ , con lo que la ecuación a estimar del tipo:

$$\ln(W_j) = \mu + \beta' \cdot X_j + \varepsilon_j$$

ecuación en la que el sesgo de habilidad ha desaparecido.

Antes de estimar los retornos a la educación con la metodología de pseudo-panel se hacen las representaciones gráficas de la evolución de los ingresos laborales y la educación. El gráfico 7.1 muestra el *stock* de capital humano por cohorte para hombres y mujeres, medido como los años medios de educación. En la cohorte de hombres se observa que la acumulación de capital humano ha crecido en forma lenta. Por ejemplo, el promedio de años de educación de la cohorte de 1945 fue alrededor de 8 años mientras que en la cohorte 1963 fue de 9 años, y la diferencia en años de escolaridad de una cohorte a la siguiente tres años más joven en promedio fue de 0,2 años.

Por su parte, la acumulación de capital humano en las mujeres se ha incrementado significativamente y en forma más rápida. Se observa que en la cohorte de 1945 los años de educación eran de 7 años y para la cohorte de 1972 los años fueron 10. La diferencia en años de la primera a la segunda cohorte mencionada fue alrededor de 0,4 años. Se nota que las mujeres han tendido a educarse más y con mayor rapidez que los hombres, aunque los hombres en las cohortes más antiguas presentan mayores niveles educativos que las mujeres en las mismas cohortes.

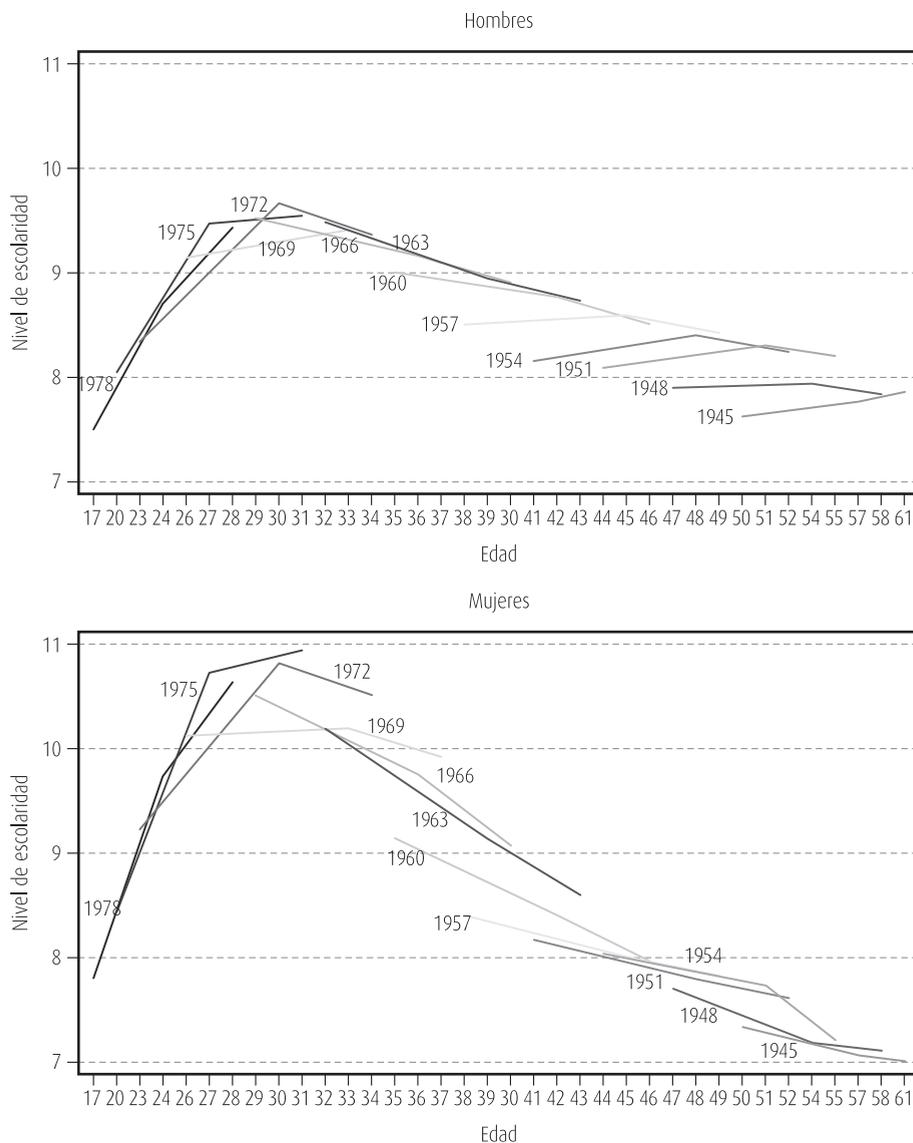
En el gráfico 7.2 se muestra el perfil de salarios para cada una de las cohortes. Para el conjunto de hombres se observa que los ingresos laborales son mayores en los individuos con mayor edad y hay una diferencia promedio de 0.7 de una cohorte a otra. Por su parte en las mujeres se observa que en las primeras tres cohortes (1945, 1948 y 1951) los salarios no han variado mucho, esto posiblemente debido a los bajos niveles de educación que se observan en las cohortes más antiguas (véase gráfico 7.1).

Ahora se pasa a la estimación de los retornos a la educación. La ecuación de salarios para calcular las tasas de retorno de la educación presenta la siguiente estructura:

$$\ln W_{i(t)t} = \beta_0 + \beta_1 \cdot S_{i(t)t} + \beta_2 \cdot \text{Expe}_{i(t)t} + \beta_3 \cdot \text{Expe}_{i(t)t}^2 + \mu_{i(t)t} + \varepsilon_{i(t)t}$$

Gráfico 7.1

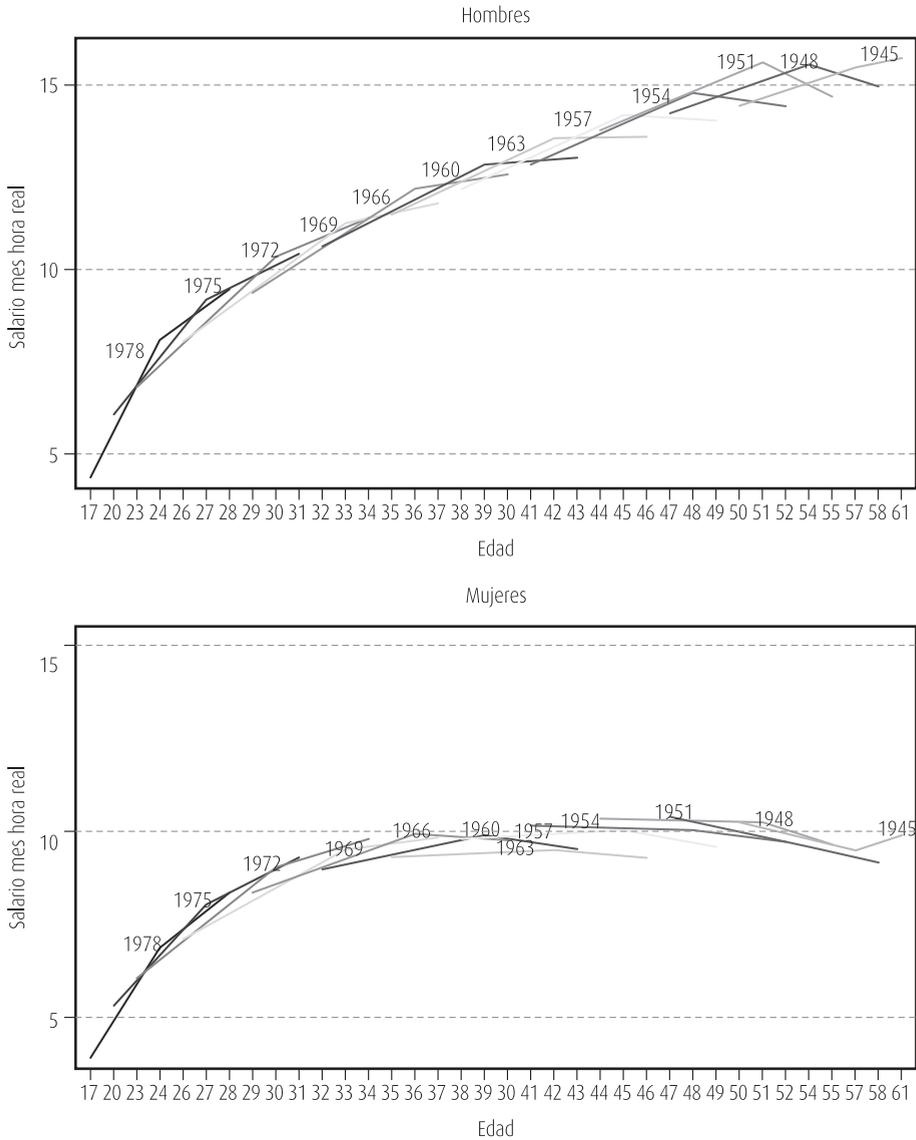
## PROMEDIO DEL NIVEL EDUCATIVO POR COHORTE Y GÉNERO



Fuente: Cálculos propios a partir de la EES.

Gráfico 7.2

## PROMEDIO DEL SALARIO REAL POR COHORTE Y GÉNERO



Fuente: Cálculos propios a partir de la EES.

En esta ecuación  $i(t)$  indica que se tienen diferentes individuos en el tiempo y  $t$  es la fecha de realización de la encuesta (1995, 2002 y 2006).  $W$  representa los salarios mensuales reales por hora;  $S$  son los años de educación;  $Expe$  es la experiencia potencial que se calcula como edad menos los años de educación menos 6;  $\mu_{i(t)}$  son efectos individuales inobservables (que incluyen, por ejemplo, la habilidad innata de los individuos) y es el término de perturbación aleatoria.

La estimación de la ecuación de salarios por MCO lleva a estimaciones inconsistentes de  $\beta_1$  si las variables explicativas se encuentran correlacionadas con  $\mu_{i(t)}$ , ya que este último es inobservable y, por tanto, se añadiría al término de perturbación. Si se dispusiera de un panel de datos genuino con un horizonte temporal muy dilatado que permitiera observar a los mismos individuos con distintos niveles educativos, este problema de inconsistencia podría resolverse utilizando el enfoque de efectos fijos. Para solventar este problema a partir de información de repetidas secciones cruzadas se tienen dos alternativas metodológicas: la de Deaton (1985) y la de Moffitt (1993).

La propuesta de Deaton (1985) consiste en agregar los datos por las cohortes y luego tomar la media de todas las observaciones de la cohorte. Con lo anterior el modelo queda de la forma:

$$\overline{\ln W_{ct}} = \beta_0 + \beta_1 \cdot \overline{S_{ct}} + \beta_2 \cdot \overline{Expe_{ct}} + \beta_3 \cdot \overline{Expe_{ct}^2} + \overline{\mu_{ct}} + \overline{\varepsilon_{ct}}$$

Entonces, por ejemplo,  $\overline{S_{ct}}$  es el valor promedio de todos los valores de  $S$  en la cohorte  $c$  en el tiempo  $t$ . El problema con la estimación de la anterior ecuación es que  $\overline{\mu_{ct}}$  depende de  $t$ , es inobservable y está probablemente correlacionado con las medias de las variables explicativas (si  $\mu_{i(t)}$  está correlacionado con las variables explicativas). Por lo tanto, si se trata  $\overline{\mu_{ct}}$  como un error aleatorio llevará a estimadores inconsistentes. Alternativamente, se puede tratar  $\overline{\mu_{ct}}$  como un parámetro desconocido fijo asumiendo que las variaciones en el tiempo pueden ser ignoradas (es decir  $\overline{\mu_{ct}} = \mu_c$ ). Deaton (1985, 116) plantea que si el tamaño de las cohortes muestrales es grande, este supuesto parece razonable, lo cual se cumple para nuestros datos como se observa en la tabla 7.2.

Se tiene, entonces, que el estimador de Deaton incorpora efectos fijos de cohorte que dan cuenta de las características inobservables de los individuos y, por tanto, se evita el potencial problema de sesgo de habilidad. Tal estimador es análogo al estimador Within que se obtiene en un panel original. Adicionalmente, si el número de observaciones es muy diferente por cohorte es necesario ponderar el modelo por la raíz cuadrada del número de observaciones de cada cohorte.

Como extensión al estimador Within de Deaton, Moffitt (1993) propone un método de estimación que hace uso de los microdatos y se basa en la idea de que el agrupamiento puede ser visto como un procedimiento de variables instrumentales. Moffitt empieza descomponiendo el efecto individual  $\mu_{i(t)}$  en un efecto cohorte  $\mu_c$  más una desviación indivi-

dual de ese efecto,  $\mathbf{v}_{i(t)}$ . Definiendo  $Z_{ci(t)} = 1$  ( $c = 1, \dots, C$ ) si el individuo  $i(t)$  es un miembro de la cohorte  $c$  y 0 en otro caso, se puede escribir:

$$\mu_{i(t)} = \sum_{c=1}^C \mu_c Z_{ci(t)} + \mathbf{v}_{i(t)}$$

Sustituyendo la anterior ecuación en la ecuación de salarios, se tiene:

$$\ln W_{i(t)} = \beta_0 + \beta_1 \cdot S_{i(t)} + \beta_2 \cdot \text{Exp}e_{i(t)} + \beta_3 \cdot \text{Exp}e_{i(t)}^2 + \sum_{c=1}^C \mu_c Z_{ci(t)} + \mathbf{v}_{i(t)} + \varepsilon_{i(t)}$$

Aplicar MCO a la anterior ecuación puede conducir a estimaciones inconsistentes, si  $\mu_{i(t)}$  y las variables explicativas están correlacionados, ya que esto comportaría que  $\mathbf{v}_{i(t)}$  también estuviera correlacionado con dichas variables. Si se encontraran instrumentos para las variables explicativas que no estuviesen correlacionados con  $\mathbf{v}_{i(t)} + \varepsilon_{i(t)}$ , se podrían obtener estimadores consistentes de  $\beta_1$  y  $\mu_c$ . Verbeek (2008) propone los siguientes instrumentos para las variables explicativas y deriva el siguiente predictor lineal:

$$X_{i(t)} = \sum_{c=1}^C \sum_{t=1}^T \gamma_{1c} Z_{ci(t)} D_t + \sum_{c=1}^C \gamma_{2c} Z_{ci(t)} + \mathbf{v}_{i(t)}$$

donde  $D_t$  es una variable dummy de años y  $\mathbf{v}_{i(t)}$  es un término de error. El resultado es un estimador idéntico al estimador Within estándar de Deaton.

Las estimaciones de la ecuación de salarios por las metodologías propuestas se muestran en la tabla 7.3. Se observa que las estimaciones son estadísticamente significativas y los signos son los esperados: positivo para la educación y la experiencia, lo que muestra que las remuneraciones laborales responden positivamente ante mayor acumulación de capital humano; y negativo para la experiencia al cuadrado, que indica que después de cierta edad los conocimientos se vuelven obsoletos y no implican mayores ingresos. Se nota que las tasas de retorno estimadas son muy similares entre las metodologías, sobre todo la de Moffitt con la de Deaton ponderada, siendo esta última más adecuada que la de Deaton simple ya que corrige por heteroscedasticidad.

Se tiene entonces que, en los hombres, cada año adicional de educación incrementa las remuneraciones laborales en un 11%, mientras que en las mujeres el incremento es de 10%. Estos resultados contrastan con lo encontrado en los gráficos, donde se observa que a pesar de que las mujeres cada vez se han ido educando más, (de hecho, en generaciones recientes, las mujeres son más educadas que los hombres), los rendimientos de dicha educación de las mujeres son menores a la de los hombres. Evaluar el diferencial de rendimientos es difícil si bien una potencial explicación es que la oferta de educados ha crecido más rápido en el caso de las mujeres que en el de los hombres, lo que puede afectar negativamente al rendimiento.

Cuando se comparan las estimaciones de los rendimientos de la educación y la experiencia bajo las técnicas de pseudos-panel con las obtenidas por MCO, se observa que las primeras son mayores que las segundas. Esto puede implicar que el sesgo de habilidad tiene poca importancia cuando se estiman los rendimientos educativos por MCO.

Tabla 7.3

## ESTIMACIONES DE LA ECUACIÓN DE SALARIOS CON METODOLOGÍA DE PSEUDOPANEL

	MUESTRA DE HOMBRES			
	DEATON SIN PONDERAR	DEATON PONDERANDO	MOFFITT	MCO
<i>S</i>	0,153 (6,76)	0,114 (6,15)	0,116 (36,32)	0,083 (327,25)
<i>Expe</i>	0,056 (10,48)	0,051 (15,92)	0,052 (77,63)	0,044 (121,36)
<i>Expe</i> <sup>2</sup>	-0,0007 (-8,00)	-0,0006 (-11,13)	-0,0007 (-53,07)	-0,0004 (-66,37)
<i>N</i>	36	36	243.223	243.223
<i>Test</i> $\hat{\mu}_c = 0$	F(11, 21) = 2,34 Prob > F = 0,045	F(11, 21) = 3,04 Prob > F = 0,014	chi2(11) = 1030,3 Prob > chi2 = 0,000	
	MUESTRA DE MUJERES			
	DEATON SIN PONDERAR	DEATON PONDERANDO	MOFFITT	MCO
<i>S</i>	0,134 (11,26)	0,105 (11,99)	0,107 (38,22)	0,075 (193,26)
<i>Expe</i>	0,046 (10,95)	0,046 (18,78)	0,047 (51,27)	0,037 (76,09)
<i>Expe</i> <sup>2</sup>	-0,0006 (-8,54)	-0,0007 (-13,01)	-0,0007 (-36,75)	-0,0005 (-47,62)
<i>N</i>	36	36	105.361	105.361
<i>Test</i> $\hat{\mu}_c = 0$	F(11, 21) = 3,56 Prob > F = 0,006	F(11, 21) = 13,38 Prob > F = 0,000	chi2(11) = 724,03 Prob > chi2 = 0,000	

Fuente: Cálculos propios a partir de las EES.

( ) t estadístico.

Nota: en el anexo 2 se encuentran las estimaciones de la primera etapa de las variables instrumentales. Se han instrumentando la educación, la experiencia potencial y su cuadrado. En la estimación MCO se incluyeron variables binarias de años.

#### ■ 4.8. EL BACKGROUND FAMILIAR Y SESGO DE ENDOGENEIDAD: APLICACIÓN DE VARIABLES INSTRUMENTALES

La Encuesta de Condiciones de Vida (ECV) dispone de información de salarios y dotación de capital humano para los ocupados haciendo posible la estimación de los rendimientos educativos. La ECV desarrolla cada año un módulo que permite profundizar sobre algunos aspectos. En el año 2005, incluyó el módulo de "Transmisión intergeneracional de la pobreza" que contiene información sobre el nivel educativo de los padres. Esta variable tiene una elevada importancia en el campo académico, pues permite controlar el sesgo de endogeneidad en la estimación de los rendimientos educativos. Es decir, los hijos cuyos padres tienen mayor nivel educativo posiblemente gozarán de un entorno que les facilitará la consecución de mayores logros educativos.

La situación ideal a efectos de corregir el denominado sesgo de endogeneidad se daría si los padres educados transmitiesen a sus hijos sus preferencias por la educación, pero no les transmitiesen su habilidad innata. Una muestra de este tipo la ofrecería disponer de información sobre niveles educativos de los padres y niveles educativos de los hijos adoptivos. En este caso, por definición, los padres podrían transmitir a sus hijos adoptivos sus prefe-

rencias por la educación, pero no podría existir transmisión genética de habilidades. La muestra utilizada no permite esta separación y el *background* familiar se emplea como instrumento bajo la hipótesis de que la transmisión intergeneracional del talento innato es escasa y que lo que muestra la correlación observada entre educación de los padres y educación de los hijos responde sólo a transmisión de la preferencia por educarse, o al mayor empeño que los padres educados ponen en que sus hijos se eduquen.

Si características como el *background* familiar, entre otros factores, afectaran tanto el nivel como los resultados educativos que el individuo obtiene, la especificación de la ecuación de salarios vendría dada por la siguiente relación:

$$(1) \quad \ln W_i = \gamma'X + \delta \cdot Z_i + \varepsilon_i$$

donde,  $\gamma$  representaría los coeficientes estimados para la dotación de capital humano en el sentido convencional (escolaridad y experiencia), y  $\delta$  recogería el efecto que tienen otras características sobre los salarios, por ejemplo el *background* familiar (vector  $Z$ ).  $\varepsilon$  es la perturbación aleatoria.

El vector de variables  $Z$  recoge un conjunto de características no observables para el individuo que normalmente son difíciles de controlar empíricamente, quedando recogidas por la perturbación aleatoria. El modelo que se estima es por tanto:

$$(2) \quad \begin{aligned} \ln W &= \gamma'X + u \\ u &= \delta \cdot Z + \varepsilon \end{aligned}$$

Los coeficientes estimados en la ecuación (2) podrían estar sesgados si existe una correlación entre el nivel de escolaridad y las variables omitidas en el modelo:  $Cov(X_i, u_i) \neq 0$ . Si poblacionalmente  $\delta \neq 0$ ,  $u_i$  estará correlacionada con la escolaridad por habilidad omitida, o por factores como la calidad de la educación o el *background* familiar.

La estimación por variables instrumentales se propone como una solución al sesgo de endogeneidad. La idea es encontrar instrumentos que estén correlacionados con la escolaridad pero que sean independientes del resto de variables omitidas. Esta alternativa, no obstante, no está exenta de problemas, pues su fiabilidad como corrección está sujeta a la validez de los instrumentos empleados (Wooldridge, 2002).

En este estudio, se propone evaluar la presencia de sesgo de endogeneidad implícito en la función de salarios, empleando como instrumento el nivel educativo del padre. Se empleará la información provista en el módulo de transmisión intergeneracional de la pobreza de la ECV-2005. Se estiman los retornos para la población de asalariados, mayores de 25 años, que han concluido su etapa formativa.

Los retornos se estimaron por medio de la especificación continua y discreta de la variable escolaridad. Para el caso discreto, se consideraron 3 categorías:

- Básicos: Hasta secundaria obligatoria y formación profesional que no requiere titulación secundaria obligatoria.

- Medios: Secundaria posobligatoria, formación profesional que requiere titulación secundaria obligatoria y diplomados.
- Superiores: Licenciados.

Para la estimación de la ecuación de salarios por niveles, se tomó el nivel de estudios básicos como referencia. La estimación de estas ecuaciones por MCO y por Variables Instrumentales se presenta en las tablas 8.1 y 8.2.

Los retornos estimados por ambas metodologías son significativos y presentan los signos esperados. Los retornos estimados a partir de la ecuación continua de salarios para el año 2005 con información de la ECV, presentan una estimación de 6,9% por MCO, frente a al 9.1% que se obtiene al aplicar variables instrumentales tomando los años de escolaridad del padre como instrumento.

Tabla 8.1

### ESTIMACIÓN DE LA ECUACIÓN DE SALARIOS CONTINUA

MÍNIMOS CUADRADOS ORDINARIOS		
VARIABLE	COEFICIENTE	ESTADÍSTICO T
Constante	0,7997	23,0
Años escolaridad	0,0687	44,1
Años experiencia	0,0323	16,1
Años experiencia <sup>2</sup>	-0,00035	-9,1
Mujer	-0,2083	-17,6
R Cuadrado	31%	
Error estándar	0,443	

Variable dependiente: logaritmo del salario mensual por hora. Método: Mínimos Cuadrados Ordinarios con Factores de Elevación Poblacionales y errores estándar robustos por White. Muestra: 7.889 observaciones.

VARIABLES INSTRUMENTALES		
VARIABLE	COEFICIENTE	ESTADÍSTICO T
Constante	0,544	8,9
Años escolaridad	0,0906	22,9
Años experiencia	0,031	11,5
Años experiencia <sup>2</sup>	-0,00028	-5,2
Mujer	-0,2324	-18,5
R Cuadrado	27%	
Error estándar	0,454	

Variable dependiente: Logaritmo del salario mensual por hora. Método: Variables Instrumentales con Factores de Elevación Poblacionales y errores estándar robustos por White. Instrumentos: Constante, Años de escolaridad del padre, Edad, Edad<sup>2</sup>, Sexo. Muestra: 7877 Observaciones.

Fuente: Cálculos propios a partir de las ECV-2005.

Tabla 8.2

### ESTIMACIÓN DE LA ECUACIÓN DE SALARIOS POR NIVELES EDUCATIVOS

MÍNIMOS CUADRADOS ORDINARIOS		
VARIABLE	COEFICIENTE	ESTADÍSTICO T
Constante	1,2318	44,8
Estudios Medios	0,2753	17,9
Estudios Superiores	0,7305	45,9
Años experiencia	0,0389	19,2
Años experiencia <sup>2</sup>	-0,00051	-13,4
Mujer	-0,2092	-18,0
R Cuadrado	32%	
Error estándar	0,439	

Variable dependiente: logaritmo del salario mensual por hora. Método: Mínimos Cuadrados Ordinarios con Factores de Elevación Poblacionales y errores estándar robustos por White. Muestra: 7.889 observaciones.

VARIABLES INSTRUMENTALES		
VARIABLE	COEFICIENTE	ESTADÍSTICO T
Constante	1,1729	24,8
Estudios Medios	0,3812	2,6
Estudios Superiores	0,9505	23,4
Años experiencia	0,0331	11,2
Años experiencia <sup>2</sup>	-0,00036	-5,8
Mujer	-0,2327	-18,4
R Cuadrado	27%	
Error estándar	0,454	

Variable dependiente: Logaritmo del salario mensual por hora. Método: Variables Instrumentales con Factores de Elevación Poblacionales y errores estándar robustos por White. Instrumentos: Constante, Años de escolaridad del padre, Edad, Edad<sup>2</sup>, Sexo. Muestra: 7.877 Observaciones.

Fuente: Cálculos propios a partir de las ECV-2005.

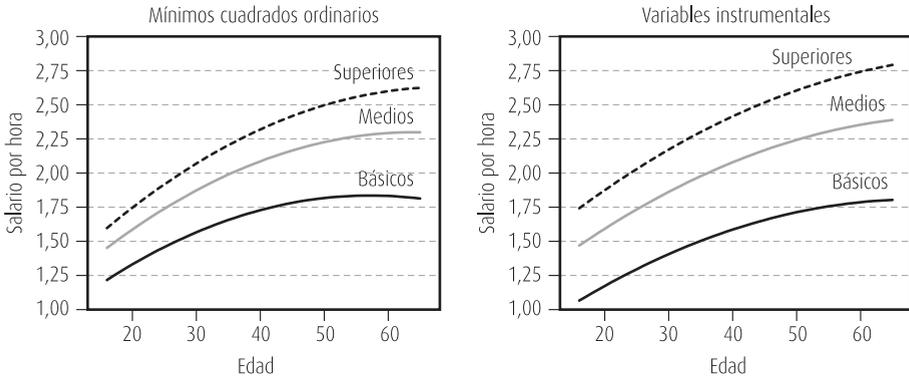
La estimación de la ecuación de salarios por niveles educativos presenta un comportamiento similar. Los retornos a la educación para la educación media y superior, respecto a los estudios básicos, es de 3,9% y 6,1%, respectivamente si se estima por MCO, y de 5,4% y 7,9%, si se hace por variables instrumentales.

Este resultado podría sorprender en el sentido, que cabría esperar que la omisión de la habilidad innata en la estimación de los retornos produjera un sesgo al alza en la medida en que el coeficiente estimado en la ecuación (2), recoja el efecto de la escolaridad y parte del efecto de la habilidad innata no observada. No obstante, los resultados presenta-

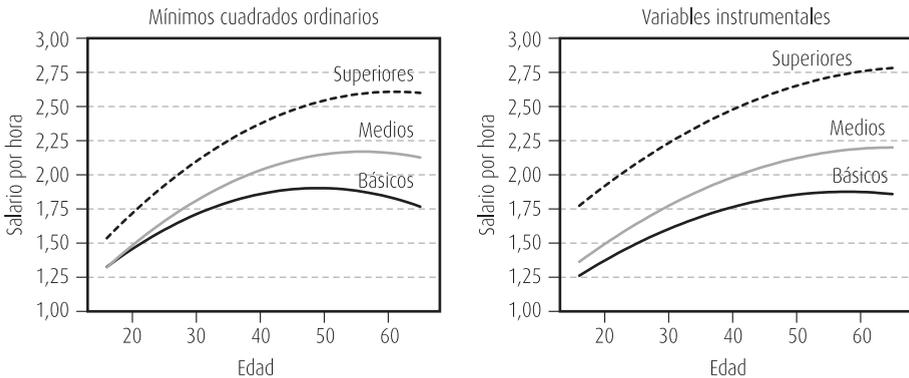
Gráfico 8.1

## PERFIL SALARIAL AL PASAR DE MCO A VI

## Estimación ecuación continua



## Estimación ecuación por niveles educativos



Fuente: Cálculos propios a partir de la ECV-2005.

dos en la estimación empleando variables instrumentales muestran una posible corrección de sesgo a la baja en dicha estimación.

Estimaciones para España empleando otras fuentes de datos y diversos instrumentos, parecen confirmar este comportamiento (Pons y Gonzalo, 2002; Arrazola *et al.*, 2003; Arrazola y Hevia, 2006; Arrazola y Hevia, 2008). La única excepción para estos resultados con datos españoles la presentan los estudios de Barceinas *et al.* (2000) y García *et al.* (2001) en los que no se aprecia variación relevante entre resultados. Parece ser que la introducción de variables instrumentales tiene como efecto un incremento de los rendimientos educativos en consonancia con los resultados de la literatura internacional.

De la estimación de los retornos por nivel educativo, se obtiene el siguiente comportamiento presentado en el gráfico 8.1 sobre el ciclo vital seguido por salarios en ambas estimaciones. El comportamiento de los perfiles salariales empleando las estimaciones continua o discreta, es muy similar.

#### ■ 4.9. DETERMINANTES DE LOS RENDIMIENTOS EDUCATIVOS

A lo largo de este estudio, hemos visto que los retornos a la educación presentan un comportamiento claramente decreciente durante el periodo 1995-2006, siguiendo sigue un patrón diferenciado según sexo, nivel educativo o región. El gráfico 9.1 presenta la correlación entre los niveles de retorno y la dotación de capital humano, o en PIB per cápita en las regiones.

Este gráfico, muestra la existencia de una relación negativa entre dotación de capital humano y físico, y los retornos a nivel regional. Las regiones que concentran la oferta de población educada o las más desarrolladas, presentan menores retornos a la educación. La oferta relativa de capital humano respecto a los requerimientos del sistema productivo es un factor fundamental en la valoración que el mercado le concede a los educados.

Con el fin de identificar los factores que condicionan el comportamiento de los rendimientos educativos, se procede a la construcción de un panel con la información de los retornos estimados a partir de la Encuesta de Estructura Salarial (EES). El panel tiene información por sexo (hombres y mujeres), para 6 niveles educativos (EGB, FP1, FP2, BUP, DIP, LIC), 17 comunidades autónomas y 3 años (1995, 2002 y 2005). En total se dispone de 612 observaciones (306 para hombres y para mujeres).

Dado que los rendimientos educativos han decrecido a la vez que ha aumentado la oferta de educados, un enfoque directo es examinar la relación existente entre rendimientos educativos, oferta de educados, las características del mercado en términos de desequilibrio entre oferta y demanda de trabajo, y la productividad implícita del empleo regional.

Estas relaciones se aproximarán a través del nivel educativo promedio de la población activa, la tasa de desempleo según nivel educativo y la proporción de licenciados sobre el total de ocupados, para cada región y periodo de tiempo considerado.

La ecuación ajustada es del tipo:

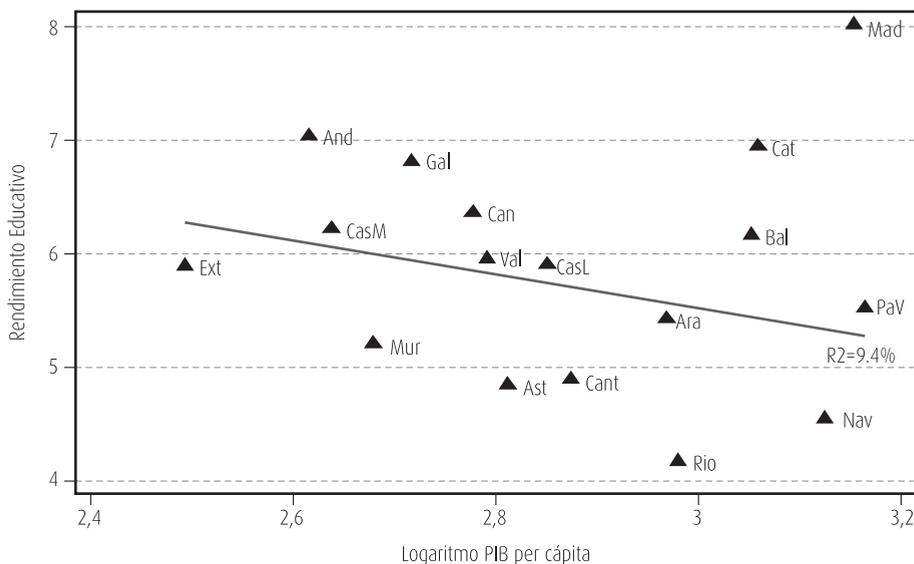
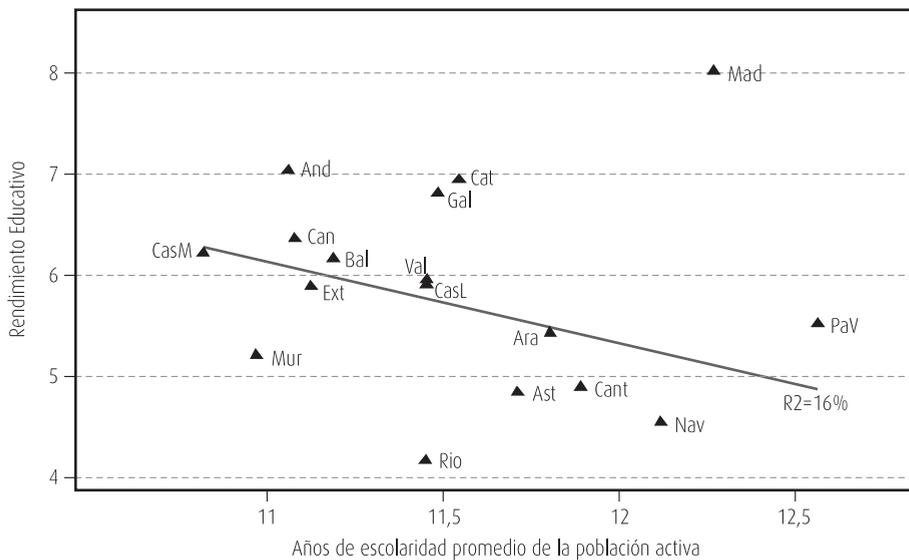
$$\text{Rendimiento}_{it} = \beta_0 + \beta_1 \cdot (\text{Años educación población activa}) + \beta_2 \cdot (\text{Tasa de desempleo})_{it} + \beta_3 \cdot \ln(\text{Proporción de ocupados licenciados sobre total})_{it} + \varepsilon_{it}$$

donde  $\varepsilon_{it}$  es la perturbación aleatoria para  $i = 1, 2, \dots, 306$  unidades cross sección observadas para  $t = 1995, 2002$  y  $2006$ .  $\beta_j$  para  $j = 0, \dots, 3$  representan los parámetros a estimar. Se espera que  $\hat{\beta}_1 < 0$ ,  $\hat{\beta}_2 < 0$  y  $\hat{\beta}_3 > 0$ .

El modelo se estima para el conjunto de la población, y por separado para hombres y mujeres. En la estimación general se incluye la *dummy* de sexo que toma valores de uno

Gráfico 9.1

## CORRELACIÓN ENTRE EL RENDIMIENTO EDUCATIVO Y NIVEL DE DESARROLLO REGIONAL



\*R2 calculado excluyendo Madrid.

Fuente: Cálculos propios a partir de la ESS. Otras fuentes del IVIE y el INE. Año 2006.

para las mujeres, para controlar las diferencias por género. Las variables incluidas como explicativas, varían por región, por nivel educativo o por sexo, quedando desvirtuada la estimación del modelo por efectos fijos. La inclusión de efectos fijos en este modelo, absorbería el efecto que se pretende captar a través de las variables explicativas distorsionando la finalidad de la estimación. Por otra parte, el número de cross sección no es lo suficientemente grande como para estimar el modelo suponiendo la presencia de efectos aleatorios (Baltagi, 2008; Wooldridge, 2002). Por tanto, se optó por la estimación de un *pooling* de datos permitiendo que la matriz de varianzas y covarianzas sea robusta a la presencia de perturbaciones no esféricas. Los resultados se presentan en la tabla 9.1.

El modelo es significativo en su conjunto y las variables presentan los signos esperados. Los años de educación de la población activa actúan con signo negativo y claramente significativo, lo que es indicativo de que en el período contemplado la oferta de educados ha superado a la demanda y ello se ha traducido en un decrecimiento de los rendimientos de la educación. En este sentido, los niveles educativos con mayor vulnerabilidad al desempleo, son también los que presentan menores tasas de rentabilidad.

No obstante, los retornos se encuentran condicionados directamente por la estructura productiva de las regiones, en cuanto a la capacidad para emplear a una mayor proporción de licenciados. Las regiones que se especializan en actividades productivas que requieren mayor dotación de licenciados tienen mayores niveles de retornos educativos.

Tabla 9.1

### DETERMINANTES DE LOS RETORNOS EDUCATIVOS PARA ESPAÑA

VARIABLE	COEFICIENTE	ESTADÍSTICO T
Constante	33,21	43,58
Años de escolaridad de los Activos	-2,737	-50,64
Tasa de Desempleo por Nivel educativo	-0,073	-10,34
Proporción de Licenciados sobre ocupados	0,380	27,77
Mujer	-0,854	-12,06
Media Variable Dependiente	10,4	
R2 Ajustado	67%	
F-Estadístico	309,0	
Número de Observaciones	612	
Periodos Incluidos	3	
Cross Sección incluidos	204	

Variable dependiente: RETORNO. Método: Panel EGLS (Cross-section weights). Muestra: 1995-2006. Panel Balanceado. Estimación White-cross sección de la matriz de varianzas y covarianzas.

Fuente: Cálculos propios a partir de la EES.

Por último, la *dummy* para mujeres resulta significativa y negativa. Las mujeres han tenido durante la última década un retorno 0,85 puntos porcentuales por debajo del retorno de los hombres.

La variable que mayor impacto tiene sobre los retornos son los años de escolaridad de la población activa (estadístico  $t = 50,6$ ). El aumento de un año en la escolaridad de la población activa, reduce en 2,7 puntos porcentuales los retornos educativos. Adicionalmente, un aumento de un punto porcentual en la proporción de licenciados dentro de los ocupados para cada región, aumenta los retornos en 0,38 puntos porcentuales.

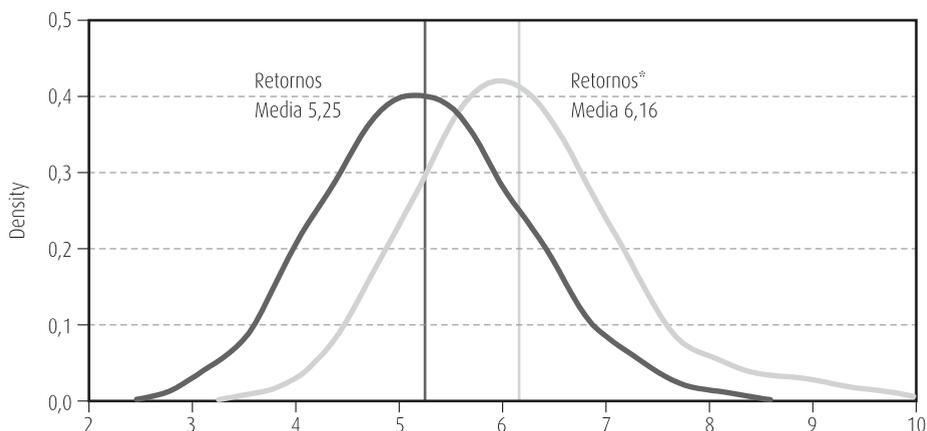
El gráfico 9.2 recoge los resultados de un ejercicio de simulación que pretende cuantificar el efecto que tendría sobre los retornos educativos, el aumento en un 20% de la proporción de licenciados sobre los ocupados para cada región. Es decir, cómo cambiaría la distribución de los retornos, si una región que tiene un 5% de proporción de licenciados, la aumenta a un 6%, o si sube de 10% a 12%.

A partir del comportamiento de los retornos para 2006, tendríamos que si la proporción de licenciados aumentara en un 20% para todas las regiones, los retornos pasarían de tener una media de 5,3% a un 6,1%.

Este ejercicio, a pesar de tener interés sólo descriptivo, por lo irreal de suponer que todas las regiones tengan la capacidad de aumentar en la misma proporción su capacidad

Gráfico 9.2

### COMPORTAMIENTO DE LOS RETORNOS FRENTE A CAMBIOS EN LA PROPORCIÓN DE LICENCIADOS SOBRE OCUPADOS



\*Retornos con aumento del 20% en la proporción de licenciados.

Fuente: Cálculos propios a partir de la ESS. 2006.

tecnológica, nos permite mostrar que más allá del impacto negativo que tienen sobre los retornos el cambio demográfico que sufre la población española que aumentó los años de escolaridad promedio, la forma directa de contrarrestar la reducción de los retornos educativos (y por ende, el desincentivo que pueden percibir los jóvenes hacia educarse), está en el cambio de la estructura productiva española, hacia actividades con mayor dotación tecnológica, con mayores niveles de productividad que demanden individuos cualificados. Esta sería una doble vía para estimular el incentivo privado de la educación, y crecer y desarrollarse económicamente.

Al estimar esta relación por separado para hombres y mujeres se obtiene:

La tabla 9.2 nos muestra que el impacto de las variables explicativas sobre los retornos para hombres y mujeres difiere. Para los hombres, las variables con mayor estadístico t son la proporción de licenciados y los años de educación de los activos (con 51.8 y 26.1, respectivamente). El aumento en un año en la escolaridad, reduce los retornos para los hombres en 3.1 puntos porcentuales, mientras el aumento de un punto porcentual en la proporción de licenciados los aumenta en 0.38 puntos porcentuales.

En el caso de las mujeres, las variable que presentan un mayor estadístico t son la tasa de desempleo por nivel educativo y los años de escolaridad de los activos (32.8 y 20.0, respectivamente). Sin embargo, el efecto marginal de la tasa de desempleo por niveles educativos no es muy alto. Un aumento en un punto porcentual en la TD por nivel educativo, reduce los retornos en 0.03 puntos porcentuales. A pesar de que el impacto no es alto, los altos niveles de desempleo que presentan las mujeres, o una mayor sensibilidad de estas al ciclo, puede repercutir en

Tabla 9.2

### DETERMINANTES DE LOS RETORNOS EDUCATIVOS SEGÚN SEXO

VARIABLE	HOMBRES		MUJERES	
	COEFICIENTE	ESTADÍSTICO T	COEFICIENTE	ESTADÍSTICO T
Constante	37,47	26,12	27,66	33,74
Años de escolaridad de los Activos	-3,097	-26,11	-2,343	-19,96
Tasa de Desempleo por Nivel educativo	-0,105	-10,11	-0,033	-32,76
Proporción de Licenciados sobre ocupados	0,385	51,80	0,378	7,03
Media Variable Dependiente		11,1		9,9
R2 Ajustado		66%		66%
F-Estadístico		201,5		201,1
Número de Observaciones		306		306
Periodos Incluidos		3		3
Cross Sección incluidos		102		102

Variable dependiente: RETORNO. Método: Panel EGLS (Cross-section weights). Muestra: 1995-2006. Panel Balanceado. Estimación White-cross sección de la matriz de varianzas y covarianzas.

Fuente: Cálculos propios a partir de la EES.

una respuesta importante de los retornos en fases recesivas. Igualmente, el aumento en un año para la escolaridad reduce en 2.3 puntos porcentuales el retorno para las mujeres. El aumento de la proporción de licenciados, los aumenta en 0.38 puntos porcentuales.

#### ■ 4.9.1. Los retornos educativos y la relación con el paro

Como se ha comentado, los retornos educativos han estado claramente afectados por el aumento de la oferta de educados. Pero, ¿cuál es su relación con el ciclo económico?, ¿Es posible que la recesión aumente el incentivo de educarse o el ciclo expansivo la reduzca? En este apartado, se pretende aportar evidencia empírica sobre esta hipótesis que empezó a esbozarse en el apartado 4.5.

Para ello, y recurriendo a la información que nos ofrece el panel de datos construido se estima una ecuación en donde los rendimientos educativos se hacen depender sólo de la oferta de educados (años medios de educación de la población activa) y de la tasa de paro. La tasa de paro agregada (como indicador de ciclo económico), en la medida en que pueda afectar a la prima salarial, puede influir sobre los rendimientos educativos. Se controla por efectos fijos individuales. Los resultados se presentan en la tabla 9.3.

Los años de educación de la población activa actúan con signo negativo y claramente significativo, lo que es indicativo de que un exceso de oferta respecto a la demanda de educados se tradujo en una disminución de los rendimientos de la educación.

La tasa de paro agregada opera con signo positivo sobre los rendimientos de la educación. Ello es indicativo de que el paro puede haber afectado más intensamente a los

Tabla 9.3

#### RELACIÓN ENTRE RENDIMIENTO EDUCATIVO Y TASAS DE DESEMPLEO

VARIABLE	COEFICIENTE	ESTADÍSTICO T
Constante	15,48	10,4
Años de escolaridad de los Activos	-0,91	-7,4
Tasa de Desempleo	0,021	1,7
R2	85%	
Media Variable dependiente	5,89	
Periodos incluidos	3	
Cross-sección incluidos	204	
Número de observaciones	612	

Variable dependiente: Retorno. Método: Panel Least Squares. Muestra: 1995-2006. Panel Balanceado. Efectos fijos cross sección.

Fuente: Cálculos propios a partir de la EES.

colectivos menos cualificados y, después de controlar por la oferta de educados, haber contribuido al aumento de la prima salarial. En cualquier caso, el efecto es relativamente poco significativo.

#### ■ 4.9.2. Efecto del desempleo sobre los rendimientos educativos corregidos por la probabilidad de ocupación

Como ya se ha señalado, los rendimientos educativos se obtienen a través de la siguiente formulación:

$$\rho_{ij}^* = \frac{(\rho_j \cdot W_j - \rho_i \cdot W_i)}{(\eta_j - \eta_i) \cdot \rho_i \cdot W_i}$$

$$\rho_{ij} = \frac{(W_j - W_i)}{(\eta_j - \eta_i) \cdot W_i}$$

$$\frac{\rho_{ij}^*}{\rho_{ij}} = \frac{(\rho_j \cdot W_j - \rho_i \cdot W_i)}{(\eta_j - \eta_i) \cdot \rho_i \cdot W_i} \cdot \frac{(\eta_j - \eta_i) \cdot W_i}{(W_j - W_i)} = \frac{(\rho_j \cdot W_j - \rho_i \cdot W_i)}{\rho_i \cdot (W_j - W_i)}$$

$$\begin{aligned} \frac{\rho_{ij}^*}{\rho_{ij}} &= \frac{\left(\rho_j \cdot \frac{W_j}{W_i} - \rho_i\right)}{\rho_i \cdot \left(\frac{W_j}{W_i} - 1\right)} = \frac{\rho_i \cdot \left(\frac{\rho_j}{\rho_i} \cdot \frac{W_j}{W_i} - 1\right)}{\rho_i \cdot \left(\frac{W_j}{W_i} - 1\right)} = \frac{\left(\frac{\rho_j}{\rho_i} \cdot \frac{W_j}{W_i} - 1\right)}{\left(\frac{W_j}{W_i} - 1\right)} = \frac{\left(\frac{\rho_j}{\rho_i} \cdot \frac{W_j}{W_i} - 1\right)}{\left(\frac{W_j - W_i}{W_i}\right)} = \left(\frac{\rho_j}{\rho_i} \cdot \frac{W_j}{W_i} - 1\right) \cdot \frac{W_i}{(W_j - W_i)} \\ \frac{\rho_{ij}^*}{\rho_{ij}} &= -\frac{W_i}{(W_j - W_i)} + \frac{\rho_j}{\rho_i} \cdot \frac{W_j}{W_i} \cdot \frac{W_i}{(W_j - W_i)} = -\frac{W_i}{(W_j - W_i)} + \frac{\rho_j}{\rho_i} \cdot \frac{W_j}{(W_j - W_i)} \end{aligned}$$

en donde  $\rho_{ij}^*$  son los rendimientos educativos considerando la probabilidad de empleo en el paso de los niveles  $i$  a  $j$ ,  $\rho_{ij}$  son los rendimientos educativos sin considerar la probabilidad de empleo,  $W_i$  son los respectivos niveles salariales y  $\rho_i$  son las respectivas probabilidades de estar ocupado obtenidas como la diferencia entre la unidad y la probabilidad de paro.

Por tanto:

$$(1) \rho_{ij}^* = \frac{W_i}{(W_j - W_i)} \cdot \rho_{ij} + \frac{W_j}{(W_j - W_i)} \cdot \frac{\rho_j}{\rho_i} \cdot \rho_{ij} = \gamma_{ij} \cdot \rho_{ij} + \delta_{ij} \cdot \rho_{ij}$$

$$\gamma_{ij} = \frac{W_i}{(W_j - W_i)}$$

$$\delta_{ij} = \frac{W_j}{(W_j - W_i)} \cdot \frac{\rho_j}{\rho_i}$$

Cabe seguidamente suponer que el desempleo afecta de forma heterogénea a los distintos niveles educativos de forma que tanto las primas salariales como las probabilidades relativas de empleo son función del valor agregado de la tasa de ocupación "to". Es decir:

$$(2) \quad \gamma_{ij} = a_1 + b_1 \cdot to + \varepsilon_1$$

$$(3) \quad \delta_{ij} = a_2 + b_2 \cdot to + \varepsilon_2$$

en donde "ε" son perturbaciones aleatorias.

Al sustituir (2) y (3) en (1) se obtiene una relación lineal del siguiente tipo:

$$\rho_{ij}^{**} = \mu \cdot \rho_{ij} + \tau \cdot (to \cdot \rho_{ij}) + u_{ij}$$

$$to = 1 - td$$

$$\rho_{ij}^{**} = \mu \cdot \rho_{ij} + \tau \cdot (1 - td) \cdot \rho_{ij} + u_{ij} = (\mu + \tau) \cdot \rho_{ij} - \tau \cdot td \cdot \rho_{ij} + u_{ij} = \alpha \cdot \rho_{ij} + \beta \cdot (td \cdot \rho_{ij}) + u_{ij}$$

Como la ecuación (1) muestra, cuando el rendimiento de la educación sin corregir es cero, el rendimiento corregido debe también ser cero. Por tanto, en la ecuación que relaciona ambos tipos de rendimientos no tiene sentido por razones conceptuales la introducción de efectos individuales. Para conocer la relación existente entre los retornos con y sin probabilidad de desempleo y el ciclo, estimamos la ecuación presentada en la tabla 9.4.

En promedio, la tasa de paro tiene un efecto negativo sobre los rendimientos de la educación. No obstante, este efecto varía por niveles. Así, al estimar la ecuación para los muy educados y los poco educados el resultado es:

La tabla 9.5 muestra que el aumento en la tasa de paro tiende a incrementar los rendimientos educativos de los más educados. Es decir, cuando la tasa de paro aumenta, este

Tabla 9.4

### RENDIMIENTOS EDUCATIVOS CON Y SIN CONTROL POR PROBABILIDAD DE DESEMPLEO Y TASA DE PARO

VARIABLE	COEFICIENTE	ESTADÍSTICO T
Retorno	1,067	66,3
Retorno * TD	-0,0054	-6,1
R2		73%
Media Variable dependiente		5,75
Periodos incluidos		3
Cross-sección incluidos		204
Número de observaciones		612

Variable Dependiente: Retorno con probabilidad de desempleo. Método: Panel Least Squares. Muestra: 1995-2006. Panel Balanceado.

Fuente: Cálculos propios a partir de la EES.

Tabla 9.5

### RENDIMIENTOS EDUCATIVOS CON Y SIN CONTROL POR PROBABILIDAD DE DESEMPLEO Y TASA DE PARO POR NIVELES EDUCATIVOS

VARIABLE	MÁS EDUCADOS*		MENOS EDUCADOS	
	COEFICIENTE	ESTADÍSTICO T	COEFICIENTE	ESTADÍSTICO T
RETORNO	1,004	79,4	1,100	48,7
RETORNO * TD	0,0034	4,7	-0,0107	-8,7
R2	84%		70%	
Media Variable dependiente	7,25		5,00	
Periodos incluidos	3		3	
Cross-sección incluidos	68		136	
Número de observaciones	204		408	

Variable dependiente: Retorno con probabilidad de desempleo. Método: Panel Least Squares. Muestra: 1995-2006. Panel Balanceado. \*Diplomados y licenciados.

Fuente: Cálculos propios a partir de la EES.

aumento suele reducir el coste de oportunidad de educarse lo que incentiva la entrada en la formación superior.

El mismo resultado se obtiene al calcular una simple diferencia de rendimientos por niveles educativos (tabla 9.6):

Para concluir los retornos educativos en España durante el periodo 1995-2006, se han visto afectados negativamente por el desequilibrio entre la oferta y demanda de educados. Un aumento de la oferta de educados en una proporción mayor a la que el mercado puede absorber, reduce la rentabilidad de la educación.

Tabla 9.6

### EFFECTO DE LA TASA DE PARO SOBRE EL DIFERENCIAL DE RENDIMIENTOS CON Y SIN CONTROL POR PROBABILIDAD DE DESEMPLEO Y POR NIVELES EDUCATIVOS

VARIABLE	MÁS EDUCADOS*		MENOS EDUCADOS	
	COEFICIENTE	ESTADÍSTICO T	COEFICIENTE	ESTADÍSTICO T
C	0,036	0,4	0,677	5,4
TD	0,0243	4,6	-0,0774	-10,3
Periodos incluidos	3		3	
Cross-sección incluidos	68		136	
Número de observaciones	204		408	

Variable dependiente: (Retorno con probabilidad de desempleo - Retorno). Método: Panel Least Squares. Muestra: 1995-2006. Panel Balanceado. \*Diplomados y licenciados.

Fuente: Cálculos propios a partir de la EES.

En este sentido, la estructura productiva de las regiones cumple un papel fundamental en el ajuste. Especializar la actividad económica española hacia sectores intensivos en tecnología puede ser la clave para recuperar el nivel de los salarios y generar mayores incentivos a la población para educarse, siendo éste además un camino certero hacia la senda de crecimiento.

Finalmente, se reconoce que el ciclo económico puede afectar también en el corto plazo a la decisión de educarse. Este estudio encuentra que la relación entre ciclo económico y retornos es asimétrica entre individuos más y menos cualificados. Un aumento en la tasa de desempleo, aumenta los retornos para individuos con estudios universitarios, mientras los reduce para el resto. El ciclo recesivo, aumenta la brecha entre las primas salariales entre los más y los menos educados.

#### ■ 4.10. EXTERNALIDADES DEL CAPITAL HUMANO

Una hipótesis comúnmente a contrastar en la teoría del capital humano es la existencia de externalidades de la acumulación de educación y formación. En este apartado se pretende mostrar evidencia de tales externalidades para el caso del mercado laboral español. En una primera parte se desarrollara el modelo analítico y en la segunda se muestran las estimaciones resultantes a partir de la EES de 1995, 2002 y 2006.

##### ■ 4.10.1. Función de producción y salarios

Se parte de una función de producción del tipo Cobb-Douglas donde se relacionan los salarios con diferentes niveles de cualificación de los ocupados y el *stock* de capital, de la forma:

$$Y = A \cdot N_1^{\alpha_1} \cdot N_2^{\alpha_2} \cdot K^{\alpha_3}$$

donde  $Y$  es el nivel de producción,  $N_1$  representa los ocupados no cualificados,  $N_2$  los ocupados cualificados y  $K$  el *stock* de capital. Adicionalmente, se supone que los trabajadores cualificados generan externalidades a través de la expresión:

$$A = \Phi \cdot N_1^{\tau}$$

Por último, se supone que el salario de los trabajadores no cualificados, que no generan externalidades, se corresponde con su productividad marginal, situación que no necesariamente se da en el caso de los cualificados, ya que su productividad marginal puede exceder a su salario. Sustituyendo la anterior expresión en la función de producción, se obtiene lo siguiente:

$$Y = \Phi \cdot N_2^{\tau} \cdot N_1^{\alpha_1} \cdot N_2^{\alpha_2} \cdot K^{\alpha_3}$$

Derivando con respecto a  $N_1$  y tomando logaritmos, se tiene:

$$\frac{dY}{dN_1} = W_1 = \Phi \cdot \alpha_1 \cdot N_1^{(\alpha_1 - 1)} \cdot N_2^{(\alpha_2 + \tau)} \cdot K^{\alpha_3}$$

$$\ln W_1 = \alpha_0 - (1 - \alpha_1) \cdot \ln N_1 + (\alpha_2 + \tau) \cdot \ln N_2 + \alpha_3 \cdot \ln K$$

donde  $W_1$  representa la productividad marginal de los ocupados no cualificados, lo cual como se mencionó se puede aproximar con los salarios dada la ausencia de externalidad de estos individuos.

Las externalidades de los trabajadores cualificados y, por tanto, de una mayor acumulación de capital humano, viene determinada por el valor del coeficiente  $\tau$ . Para poder identificar dichas externalidades se supondrá que existen economías de escala constantes, es decir que  $\alpha_1 + \alpha_2 + \alpha_3 = 1$ . Entonces, a partir de la estimación de la ecuación salarial para los trabajadores no cualificados se obtendrán los coeficientes estimados y junto con la restricción de economías de escala constantes, se podrá determinar el valor de las externalidades. Esto último es lo que se pretende realizar en el siguiente apartado.

#### ■ 4.10.2. Estimaciones de las externalidades de la acumulación de capital humano

Para la estimación de la ecuación de salario de los no cualificados es necesario tener información de los niveles salariales de los no cualificados ( $W_1$ ), alguna medida de empleo cualificado ( $N_2$ ) y no cualificado ( $N_1$ ), y del *stock* de capital ( $K$ ). Para la estimación de estas variables se tuvo en cuenta lo siguiente:

- Se trabajó a nivel de sector.
- Las variables de empleo cualificado y no cualificado son representadas por el número de individuos con dichas características. Para la definición del tipo de empleo se hicieron dos agrupaciones. En una primera definición se agrupó en trabajadores cualificados a los individuos con licenciatura y en los no cualificados a los que tienen menos de licenciatura. En la segunda agrupación se definió como cualificados a la suma de trabajadores con diplomado y licenciatura, y como no cualificados al resto.
- La información del *stock* de capital se tomó del IVIE y corresponden al *stock* de capital neto real en maquinaria y equipo de transporte, como proxies del capital productivo.
- La base de datos final tendrá una estructura de panel a nivel de 21 sectores en 3 años.

Para el cálculo de los niveles salariales de los no cualificados ( $W_1$ ) se estimó la siguiente ecuación en cada uno de los años de la EES:

$$\ln(\text{salario}_i) = \beta_1 \text{exper} + \beta_2 \text{exper}_i^2 + \beta_3 \text{mujer}_i + \underbrace{\sum_{s=1}^{21} \gamma_s \cdot F_s}_{\text{Dummies de sector}} + u_i$$

donde *salario* es el salario real hora, *exper* es la experiencia potencial, *mujer* es una variable binaria que toma el valor de 1 para mujer y 0 para hombre,  $F_s$  representa variables binarias que indican 1 cuando es el sector  $s$  y 0 otro sector, y  $u$  es el término de perturbación aleatoria. Se tiene entonces, que los coeficientes  $\gamma_s$  miden el logaritmo salarial sectorial para los no cualificados después de descontar el efecto derivado de la experiencia y el género (en el anexo 3 se muestran los resultados de las estimaciones). Los resultados obtenidos fueron los siguientes (tabla 10.1).

Las estimaciones muestran una elasticidad de los salarios negativa respecto a los trabajadores no cualificados y positiva respecto a los cualificados y al *stock* de capital. Lo anterior indica que a mayor empleo cualificado y capital físico, la productividad, por medio de los salarios, será mayor en los empleados menos cualificados. Existe, por tanto, una complementariedad y efectos externos positivos de una mayor dotación de capital humano y físico. Por el contrario, un crecimiento de la mano de obra menos cualificada disminuye su propia productividad, debido a la falta de externalidades de este tipo de ocupados.

A partir de los anteriores resultados es posible calcular el valor de las externalidades que generan los ocupados más cualificados, es decir el valor de de la función de produc-

Tabla 10.1

### ESTIMACIONES DE LA FUNCIÓN DE PRODUCCIÓN Y SALARIOS ( $Y = \ln(W_1)$ )

	LICENCIADOS	LICENCIADOS + DIPLOMADOS
$\ln(N_1)$	-0,274 (-7,92)	-0,272 (-6,76)
$\ln(N_2)$	0,192 (3,75)	0,159 (2,52)
$\ln(K)$	0,209 (5,80)	0,231 (6,64)
Efectos fijos de sector ( $\hat{\alpha}_s$ )	Sí	Sí
N	63	63
R <sup>2</sup>	0,956	0,956
Ho: $\hat{\alpha}_s = 0$	F(20, 39) = 5,63; Prob > F = 0,000	F(20, 39) = 5,29; Prob > F = 0,000
Test de Hausman	Chiq(3) = 16,78	Chiq(3) = 24,17
Ho: Cor( $x, \hat{\alpha}_s$ ) = 0	Prob > Chiq = 0,000	Prob > Chiq = 0,000

Fuente: Cálculos propios a partir de la EES.

ción y salarios. Tomando las estimaciones para la ecuación de licenciados, se obtiene lo siguiente:

$$\begin{aligned}\ln W_1 &= \alpha_0 - (1 - \alpha_1) \cdot \ln N_1 + (\alpha_2 + \tau) \cdot \ln N_2 + \alpha_3 \cdot \ln K \\ &\quad - (1 - \hat{\alpha}_1) = -0,0274 \\ \hat{\alpha}_2 + \hat{\tau} &= 0,192 \\ \hat{\alpha}_3 &= 0,298\end{aligned}$$

y junto con el supuesto de economías a escala constantes  $\hat{\alpha}_1 + \hat{\alpha}_2 + \hat{\alpha}_3 = 1$  se tiene que:

$$\begin{aligned}\hat{\alpha}_1 &= 0,726 \\ \hat{\alpha}_3 &= 0,209 \\ \hat{\alpha}_1 + \hat{\alpha}_2 + \hat{\alpha}_3 = 1 &\rightarrow 0,726 + \hat{\alpha}_2 + 0,209 = 1 \rightarrow \hat{\alpha}_2 = 0,065 \\ \hat{\tau} &= 0,127\end{aligned}$$

Para el caso en el que se toma como ocupados cualificados los que tienen diplomatura y licenciatura, se obtiene una externalidad de 0,118. Se tiene, entonces, que aparecen externalidades positivas de los ocupados con una mayor dotación de capital humano, implicando mayores productividades no sólo para los propios ocupados cualificados, sino también para los no cualificados.

Ahora bien, para probar que las externalidades de los ocupados cualificados son estadísticamente significativas se reescribe la ecuación a estimar. En efecto, partiendo del supuesto de economías de escala constantes, se verifica:

$$\alpha_3 = 1 - \alpha_1 - \alpha_2$$

Sustituyendo esta expresión en la función de producción y salarios se obtiene:

$$\begin{aligned}\ln W_1 &= \alpha_0 - (1 - \alpha_1) \cdot \ln N_1 + (\alpha_2 + \tau) \cdot \ln N_2 + (1 - \alpha_1 - \alpha_2) \cdot \ln K \\ \ln W_1 &= \alpha_0 - (1 - \alpha_1) \cdot \ln N_1 + (\alpha_2 + \tau) \cdot \ln N_2 + (1 - \alpha_1) \cdot \ln K - \alpha_2 \cdot \ln K \\ \ln W_1 &= \alpha_0 - (1 - \alpha_1) \cdot \ln \left( \frac{N_1}{K} \right) - \alpha_2 \cdot \ln \left( \frac{N_2}{K} \right) + \tau \cdot \ln N_2\end{aligned}$$

Estimando la anterior ecuación y calculando el t-estadístico convencional se prueba la significancia estadística de  $\tau$ . Los resultados de lo anterior se muestran en la tabla 10.2.

Con el anterior procedimiento se puede estimar directamente los coeficientes de la función de producción y salarios, así como sus respectivos t-estadísticos. Se puede observar que las externalidades positivas que generan los ocupados cualificados sobre la productividad de los no cualificados son estadísticamente significativas. Se corrobora, entonces, la existencia de externalidades en la acumulación de capital humano en el mercado laboral español.

Tabla 10.2

ESTIMACIONES DE LA FUNCIÓN DE PRODUCCIÓN Y SALARIOS TRANSFORMADA ( $Y = \text{Ln}(W_1)$ )

	LICENCIADOS	LICENCIADOS + DIPLOMADOS
$\text{Ln}(N_1/k)$	-0,274 (-7,92)	-0,272 (-6,76)
$\text{Ln}(N_2/k)$	0,065 (1,05)	0,041 (0,607)
$\text{Ln}(N_3)$	0,127 (2,97)	0,118 (2,83)

Fuente: Cálculos propios a partir de la EES.

#### ■ 4.11. CONCLUSIONES

Las principales conclusiones de este estudio, se recogen en los siguientes puntos:

- La dotación del capital humano para España en la última década ha aumentado, presentando importantes diferencias por regiones y según sexo:
  - España ha tenido un continuo aumento del nivel de capital humano productivo: mientras que en 1995 los años medios de educación de los ocupados fueron 10, para 2006 este indicador se ubica alrededor de 13 años medios.
  - A nivel de comunidades autónomas el anterior crecimiento se mantiene y sobrepasan Madrid, País Vasco y Navarra como las regiones con mayores dotaciones de capital humano productivo, con más de 12 años medios de educación en sus ocupados. En cuanto a la distribución geográfica del capital humano productivo se observa dos áreas bien definidas. La primera es el noroeste con mayor número de años medios de educación de sus ocupados y la segunda es el suroeste con menores niveles. Estos resultados son consistentes con la evidencia empírica encontrada en otros estudios (véase capítulo 2).
  - La acumulación de capital humano de las mujeres españolas se ha incrementado significativamente y en forma más rápida en comparación con los hombres. Se ha observado que entre las cohortes de mujeres con mayor edad existe en promedio una diferencia de tres años con respecto a las más jóvenes, mientras que en los hombres dicha diferencia sólo es de un año.
- Los retornos a la educación han caído en la última década, y ciertamente la recomposición los asalariados en términos de sus características ha influido sobre este comportamiento:
- Los rendimientos por años adicionales de educación son menores para las mujeres que para los hombres, y su caída es también más pronunciada.
- Los niveles educativos superiores han sido los más afectados por la disminución en la rentabilidad de la educación.

- El aumento en la oferta de educados, que no ha tenido su contrapartida en un aumento de la demanda, se interpreta como uno de los condicionantes de esta caída en la rentabilidad de la educación.
- En la estimación de los rendimientos educativos es siempre un tema polémico el relativo a la adecuada corrección del denominado sesgo de endogeneidad. El sesgo de endogeneidad procede de dos fuentes: En primer lugar, los más hábiles pueden mostrar una mayor propensión a educarse, y en la medida en que la habilidad innata es inobservable, la estimación de rendimientos educativos a partir de la aplicación de Mínimos Cuadrados Ordinarios en ecuaciones que excluyen la habilidad como variable explicativa, puede presentar un sesgo al alza. En segundo lugar, los salarios estimados y la decisión de educarse pueden ser dos variables que se determinen de forma conjunta. Dos alternativas de estimación se han utilizado para paliar estos sesgos. Una de ellas es la utilización de un pseudo panel, y la otra, es la aplicación de Variables Instrumentales. No obstante, al pasar de Mínimos Cuadrados Ordinarios a técnicas de pseudo panel, o al pasar de Mínimos Cuadrados Ordinarios a Variables Instrumentales se halla un incremento en lugar de una reducción en los rendimientos educativos. Este resultado es común en la literatura pero no por ello deja de ser un tanto sorprendente. Diversas explicaciones se han ofrecido al respecto. Entre ellas, la que nos parece más creíble, es que la estimación de los rendimientos educativos por Variables Instrumentales no aproxima el rendimiento promedio de la inversión en educación, sino el rendimiento de los más hábiles en un contexto en que los rendimientos educativos individuales son específicos de cada individuo. En cualquier caso, la evidencia hallada sugiere que hay razones objetivas para afirmar que los rendimientos de la educación hallados a través de la aplicación de Mínimos Cuadrados Ordinarios constituyan una sobrevaloración de los verdaderos rendimientos derivados de incrementar los años de educación.
- Se ha hallado cierta evidencia a favor de la presencia de externalidades positivas en la acumulación de capital humano. En concreto se ha estimado que la acumulación de capital humano cualificado aumenta la productividad de la mano de obra no cualificada en una proporción superior a la que cabría esperar si no existiesen externalidades positivas. En este contexto, cabe destacar la importancia de fortalecer la estructura productiva española consiguiendo una mayor orientación hacia actividades con un mayor contenido tecnológico.

## ■ REFERENCIAS

- ANTMAN, F. y MCKENZIE, D. (2007), "Earnings Mobility and Measurement Error: A Pseudo-Panel Approach", *Economic Development and Culture Change*, 56 (1): 125-161.
- ARRAZOLA, M. y J. DE HEVIA (2008), "Three measures of returns to education: an illustration for the case of Spain", *Economics of Education Review*, 27: 266-275.
- (2006), "Gender differentials in returns to education in Spain", *Education Economics*, 14, 4, 469-486.
- ARRAZOLA, M., DE HEVIA, J. y SANZ, J. F. (2003) "Returns to education in Spain: some evidence on the endogeneity of schooling", *Education Economics*, 11 (3): 293-304.
- ATANAZIO, O. (1997), "Cohort analysis of saving behavior by U.S. households", *Journal of Human Resources*, 33 (3): 576-609.
- BALTAGI, B. (2008), *Econometric Analysis of Panel Data*. 4th edition. Wiley press.
- BARCEINAS, F.; OLIVER, J.; RAYMOND, J. L., y ROIG, J. L. (2000), "Spain" en HARMON, C., I. WALKER y N. WESTERGAARD-NIELSEN (eds.) *Education and earnings in Europe: a cross country analysis of the return to education*. Cheltenham: Edward Elgar.
- BEHRMAN, J. R. y ROSENZWEIG, M. R. (1999), "Ability Biases in Schooling Returns and Twins: a Test and New Estimates", *Economics of Education Review*, 18: 159-167.
- BLACKBURN, M. y NEUMARK, D. (1991), "Omitted-Ability Bias and the Increase in the Return to Schooling", *Working Paper* 3693, NBER, Cambridge.
- (1995), "Are OLS Estimates of the Return to Schooling Biased Downward? Another Look", *Review of Economics and Statistics*, 77: 217-229.
- BRUNELLO, G. y MINIACI, R. (1999), "The Economic Return to Schooling for Italian Men. An Evaluation Based on Instrumental Variables", *Labour Economics*, 6: 509-519.
- BUTCHER, K. y CASE, A. (1994), "The Effect of Sibling Sex Composition on Women's Education and Earning", *Quarterly Journal of Economics*, 109: 531-563.
- DEATON, A. (1985), "Panel Data from Time Series of Cross Sections", *Journal of Econometrics*, 30: 109-126.
- GARCÍA, J.; HERNÁNDEZ, P. J. y LÓPEZ-NICOLÁS, A. (2001), "How wide is the gap?, An investigation of gender wage differences using quantile regression", *Empirical Economics*, 26: 149-167.
- MINCER, J. (1974). *Schooling, Experience and Earnings*, Columbia, University Press. NY.
- PONS, E. y GONZALO, M. T. (2002), "Returns to education in Spain: how reliable are instrumental variable estimates?", *Labour*, 16 (4): 747-770.
- SAAVEDRA, J. y VALDIVIA, M. (2000), "Household and Individual Decision-Making over the Life Cycle: A First Look at Evidence from Peruvian Cohorts", *Research Network Working Paper*, No R-425, BID, Washington.
- SERRANO, L. y PASTOR, J. (2005), *La Geografía de Capital Humano en España: Niveles Educativos de los Municipios*, Bancaja, Valencia.
- SERRANO, L.; RAYMOND, J.; ROIG, J. y PASTOR, J. (2006). *El Rendimiento del Capital Humano en España*, Bancaja, Valencia
- USITALO, R. (1999), "Return to Education in Finland", *Labor Economics*, 6: 569-580.
- VERBEEK, M (2008), "Pseudo-Panels and Repeated Cross-Sections", en: L. MÁTYÁS y P. SEVESTRE (eds.), *The Econometrics of panel data: handbook of theory and applications*, tercera edición, Springer Publishers: 369-365.
- WOOLDRIDGE, J. (2002). *Econometric Analysis of Cross Section and Panel Data*. The MIT press.





## **ANEXOS ESTADÍSTICOS**



## ■ ANEXO 1. RETORNOS A LA EDUCACIÓN PARA SIN ESTUDIOS Y ESTUDIOS PRIMARIOS

La muestra para individuos sin estudios, presenta para el año 2006 un aumento sustancial hasta el nivel del 7,1%, sin que este comportamiento sea aparentemente soportado por la realidad dado el aumento inminente de los años de escolaridad promedio para la población española en los últimos años.

La siguiente tabla presenta la composición de la muestra según nivel educativo, y en él se confirma un aumento paulatino de la participación de los niveles educativos superiores, frente a una reducción en los niveles inferiores, incluso al agrupar las categorías sin estudios y con estudios primarios:

Anexo 1

### COMPOSICIÓN DE LA MUESTRA POR NIVEL EDUCATIVO

NIVEL EDUCATIVO	1995		2002		2006	
	N.º OBS	%	N.º OBS	%	N.º OBS	%
Sin estudios	3.515	2,5%	2.013	1,3%	7.928	7,1%
Primarios	43.553	31,1%	41.198	27,6%	24.564	21,9%
Sin estudios y primarios	47.068	33,6%	43.211	28,9%	32.492	29,0%
Secundaria obligatoria	44.343	31,6%	47.105	31,6%	32.993	29,5%
Formación Profesional I	6.732	4,8%	9.455	6,3%	6.758	6,0%
Formación Profesional II	10.999	7,8%	13.242	8,9%	9.331	8,3%
Secundaria postobligatoria	16.277	11,6%	15.858	10,6%	12.338	11,0%
Diplomados	6.863	4,9%	8.381	5,6%	7.006	6,3%
Licenciatura y más	7.906	5,6%	12.036	8,1%	11.079	9,9%
<i>Total</i>	<i>140.188</i>	<i>100%</i>	<i>149.288</i>	<i>100%</i>	<i>111.997</i>	<i>100%</i>

\*Se aplicó los filtros de homogeneización de la muestra para comparar los resultados.

Fuente: Cálculos Propios a partir de la EES.

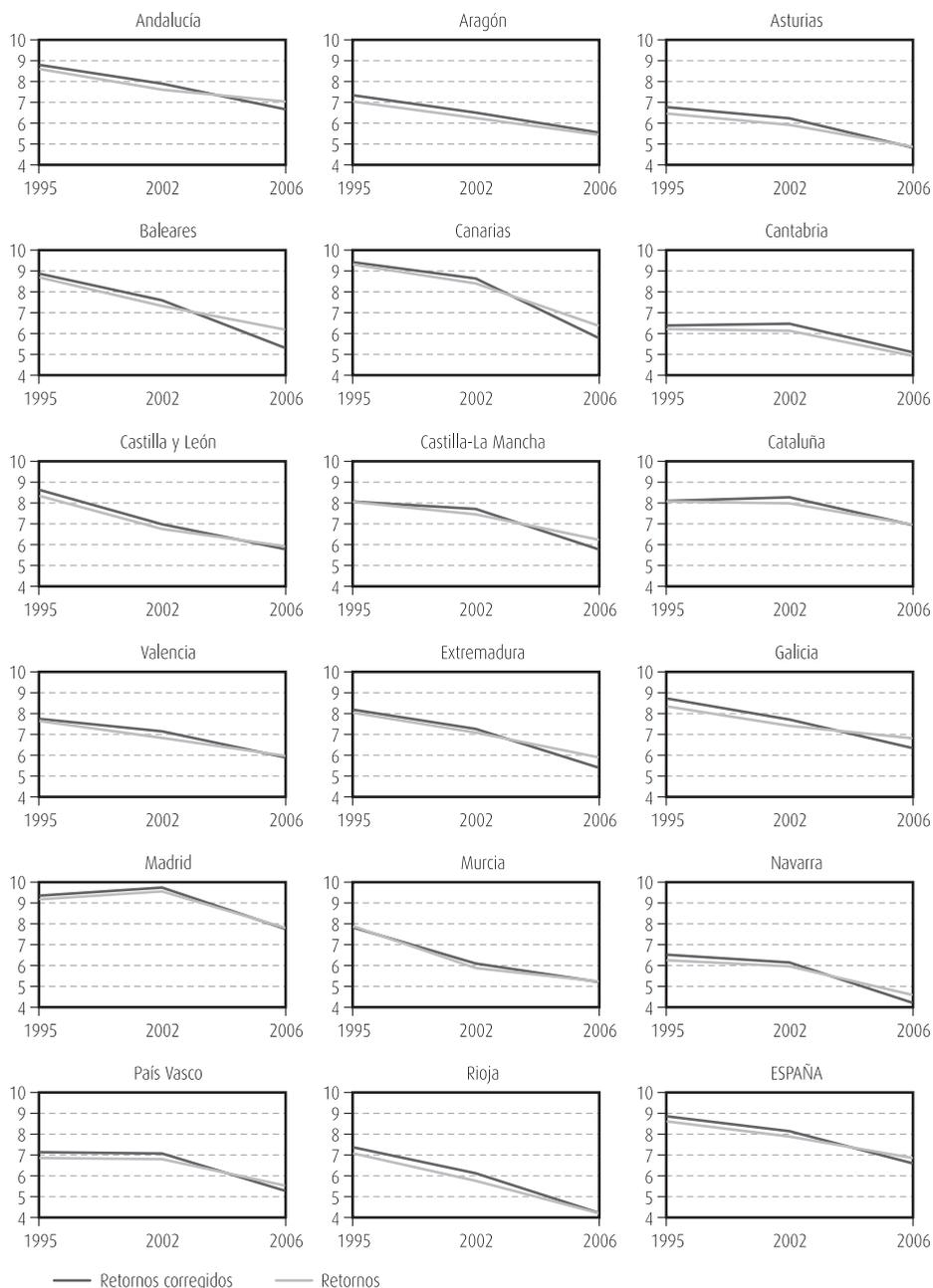
En la estimación continua de los retornos educativos se asigna el número de años correspondiente al último nivel educativo alcanzado por los individuos, por tanto, un aumento aparentemente inexplicable en la proporción de sin estudios, puede sesgar a la baja la estimación de los retornos. Es posible que el hecho que la información sobre el nivel de estudios de individuo sea aportada por el empleador, pueda conducir a algún margen de error.

Para que la comparación de los retornos educativos en el tiempo no sea afectada por posibles errores muestrales, entre otros, se procedió a agrupar a los niveles educativos de

sin estudios y con estudios primarios, asignando un número de años de estudio de 4.5 para ambos grupos en los tres periodos analizados.

El siguiente gráfico recoge el comportamiento de la serie corregida y sin corregir. Los retornos estimados aplicando la corrección (línea oscura), afecta sobre todo a las comunidades autónomas con menor tamaño muestral como por ejemplo, Baleares, Extremadura o Galicia, entre otras. No obstante, su impacto es muy leve sobre las comunidades de mayor tamaño muestral como Madrid o Cataluña. La aplicación de este ajuste es importante por tanto, para evitar posibles sesgos a la baja en la estimación de los retornos educativos para comunidades autónomas pequeñas que pueden concentrar mano de obra no cualificada.

## Gráfico



## Anexo 2

## MODELO DE MOFFITT: PRIMERA ETAPA DE LA REGRESIÓN POR VARIABLES INSTRUMENTALES

HOMBRES

Number of obs = 243.223  
 F(35,243187) = 192.67  
 Prob>F = 0.0000  
 R-squared = 0.0270  
 AdjR-squared = 0.0268  
 RootMSE = 3.7678

Y = S	COEF.	STD. ERR.	t	P >  t	[95% CONF. INTERVAL]	
cohorte2	.5747617	.1763961	3.26	0.001	.2290301	.9204933
cohorte3	.8633919	.1659778	5.20	0.000	.5380797	1.188704
cohorte4	1.663987	.1631833	10.20	0.000	1.344152	1.983822
cohorte5	2.042785	.1623236	12.58	0.000	1.724635	2.360935
cohorte6	2.002266	.1622056	12.34	0.000	1.684347	2.320185
cohorte7	1.480748	.162431	9.12	0.000	1.162388	1.799109
cohorte8	.9410287	.1626097	5.79	0.000	.622318	1.259739
cohorte9	.5976055	.1629615	3.67	0.000	.2782053	.9170057
cohorte10	.529375	.1630752	3.25	0.001	.2097518	.8489982
cohorte11	.3410476	.16289	2.09	0.036	.0217875	.6603078
cohorte12	.0918921	.1636584	0.56	0.574	-.228874	.4126583
cohor1Xaño02	1.19616	.1636229	7.31	0.000	.8754638	1.516857
cohor1Xaño06	2.027727	.1644218	12.33	0.000	1.705464	2.349989
cohor2Xaño02	1.396719	.0874043	15.98	0.000	1.225409	1.56803
cohor2Xaño06	1.617948	.0899629	17.98	0.000	1.441623	1.794273
cohor3Xaño02	1.312503	.0636478	20.62	0.000	1.187755	1.437251
cohor3Xaño06	1.186093	.0683351	17.36	0.000	1.052158	1.320028
cohor4Xaño02	.2437676	.0563638	4.32	0.000	.133296	.3542393
cohor4Xaño06	.1024774	.0620588	1.65	0.099	-.0191562	.224111
cohor5Xaño02	-.3927618	.0543891	-7.22	0.000	-.4993629	-.2861606
cohor5Xaño06	-.4717904	.0614525	-7.68	0.000	-.5922356	-.3513451
cohor6Xaño02	-.5762439	.0553548	-10.41	0.000	-.6847378	-.46775
cohor6Xaño06	-.6013318	.0619859	-9.70	0.000	-.7228225	-.4798411
cohor7Xaño02	-.2433043	.057053	-4.26	0.000	-.3551267	-.131482
cohor7Xaño06	-.3092781	.0648879	-4.77	0.000	-.4364566	-.1820996
cohor8Xaño02	.1008444	.0589039	1.71	0.087	-.0146057	.2162946
cohor8Xaño06	.1359014	.0673106	2.02	0.043	.0039743	.2678284

Anexo 2 (continuación)

### MODELO DE MOFFITT: PRIMERA ETAPA DE LA REGRESIÓN POR VARIABLES INSTRUMENTALES

Y = S	COEF.	STD. ERR.	t	P >  t	[95% CONF. INTERVAL]	
cohor9Xaño02	.2366355	.0617307	3.83	0.000	.115645	.3576259
cohor9Xaño06	.31152	.0726111	4.29	0.000	.1692041	.4538359
cohor10Xaño02	.2074766	.0636868	3.26	0.001	.0826522	.332301
cohor10Xaño06	.3006398	.076192	3.95	0.000	.1513055	.4499742
cohor11Xaño02	.0137459	.0650828	0.21	0.833	-.1138147	.1413065
cohor11Xaño06	.160762	.0811869	1.98	0.048	.0016377	.3198862
cohor12Xaño02	.089927	.0729328	1.23	0.218	-.0530194	.2328735
cohor12Xaño06	.40382	.0957202	4.22	0.000	.2162109	.5914291
_cons	7.432456	.1578149	47.10	0.000	7.123143	7.741769

Number of obs = 243223  
 F (35, 243187) = 45724.50  
 Prob > F = 0.0000  
 R-squared = 0.8681  
 Adj R-squared = 0.8681  
 Root MSE = 3.8349

Y = EXPER	COEF.	STD. ERR.	t	P >  t	[95% CONF. INTERVAL]	
cohorte2	2.282129	.1795373	12.71	0.000	1.930241	2.634018
cohorte3	4.740809	.1689335	28.06	0.000	4.409704	5.071914
cohorte4	6.898973	.1660892	41.54	0.000	6.573442	7.224503
cohorte5	9.466571	.1652143	57.30	0.000	9.142756	9.790387
cohorte6	12.48494	.1650942	75.62	0.000	12.16136	12.80852
cohorte7	15.98665	.1653235	96.70	0.000	15.66262	16.31068
cohorte8	19.47077	.1655054	117.64	0.000	19.14639	19.79516
cohorte9	22.8233	.1658634	137.60	0.000	22.49821	23.14839
cohorte10	25.86887	.1659793	155.86	0.000	25.54356	26.19419
cohorte11	29.0904	.1657907	175.46	0.000	28.76546	29.41535
cohorte12	32.38199	.1665728	194.40	0.000	32.05551	32.70847
cohor1Xaño02	5.354984	.1665367	32.15	0.000	5.028576	5.681392
cohor1Xaño06	8.590808	.1673498	51.33	0.000	8.262807	8.918809
cohor2Xaño02	5.245097	.0889608	58.96	0.000	5.070736	5.419458
cohor2Xaño06	9.162137	.0915649	100.06	0.000	8.982672	9.341602
cohor3Xaño02	5.56471	.0647812	85.90	0.000	5.437741	5.69168
cohor3Xaño06	9.855415	.069552	141.70	0.000	9.719095	9.991735
cohor4Xaño02	6.645911	.0573676	115.85	0.000	6.533472	6.75835
cohor4Xaño06	10.96409	.0631639	173.58	0.000	10.84029	11.08789

Anexo 2 (continuación)

### MODELO DE MOFFITT: PRIMERA ETAPA DE LA REGRESIÓN POR VARIABLES INSTRUMENTALES

Y = EXPER	COEF.	STD. ERR.	t	P >  t	[95% CONF. INTERVAL]	
cohor5Xaño02	7.333236	.0553576	132.47	0.000	7.224736	7.441735
cohor5Xaño06	11.60847	.0625468	185.60	0.000	11.48588	11.73106
cohor6Xaño02	7.523581	.0563405	133.54	0.000	7.413155	7.634007
cohor6Xaño06	11.72414	.0630897	185.83	0.000	11.60049	11.84779
cohor7Xaño02	7.215471	.058069	124.26	0.000	7.101657	7.329284
cohor7Xaño06	11.47629	.0660434	173.77	0.000	11.34685	11.60573
cohor8Xaño02	6.877257	.0599529	114.71	0.000	6.759751	6.994763
cohor8Xaño06	11.06507	.0685093	161.51	0.000	10.93079	11.19934
cohor9Xaño02	6.713187	.06283	106.85	0.000	6.590042	6.836332
cohor9Xaño06	10.88323	.0739042	147.26	0.000	10.73838	11.02808
cohor10Xaño02	6.762888	.0648209	104.33	0.000	6.635841	6.889935
cohor10Xaño06	10.86597	.0775488	140.12	0.000	10.71398	11.01797
cohor11Xaño02	6.890721	.0662418	104.02	0.000	6.760889	7.020553
cohor11Xaño06	10.99278	.0826327	133.03	0.000	10.83083	11.15474
cohor12Xaño02	6.777923	.0742316	91.31	0.000	6.632431	6.923415
cohor12Xaño06	10.66003	.0974248	109.42	0.000	10.46908	10.85098
_cons	4.008772	.1606252	24.96	0.000	3.693951	4.323593

Number of obs = 2243223  
 F(35, 243187) = 248125.42  
 Prob > F = 20.0000  
 R-squared = 20.8738  
 Root MSE = 2190.6486

Y = EXPER2	COEF.	STD. ERR.	t	P >  t	[95% CONF. INTERVAL]	
cohorte2	25.58795	8.925596	2.87	0.004	8.094012	43.08188
cohorte3	63.92035	8.398436	7.61	0.000	47.45963	80.38106
cohorte4	110.7323	8.257032	13.41	0.000	94.54875	126.9159
cohorte5	176.0103	8.213534	21.43	0.000	159.912	192.1086
cohorte6	267.0054	8.207564	32.53	0.000	250.9188	283.092
cohorte7	395.8438	8.218968	48.16	0.000	379.7348	411.9527
cohorte8	547.8909	8.228009	66.59	0.000	531.7642	564.0176
cohorte9	715.8837	8.245808	86.82	0.000	699.7221	732.0453
cohorte10	888.9152	8.251566	107.73	0.000	872.7424	905.0881
cohorte11	1091.839	8.242193	132.47	0.000	1075.684	1107.993
cohorte12	1320.7	8.281074	159.48	0.000	1304.47	1336.931
cohor1Xaño02	79.17877	8.279278	9.56	0.000	62.9516	95.40594
cohor1Xaño06	157.6912	8.319701	18.95	0.000	141.3848	173.9976

Anexo 2 (continuación)

### MODELO DE MOFFITT: PRIMERA ETAPA DE LA REGRESIÓN POR VARIABLES INSTRUMENTALES

Y = EXPER2	COEF.	STD. ERR.	t	P >  t	[95% CONF. INTERVAL]	
cohor2Xaño02	103.242	4.422637	23.34	0.000	94.57375	111.9103
cohor2Xaño06	213.6064	4.552098	46.92	0.000	204.6844	222.5284
cohor3Xaño02	138.4859	3.220562	43.00	0.000	132.1737	144.7981
cohor3Xaño06	283.165	3.457739	81.89	0.000	276.3879	289.942
cohor4Xaño02	194.3708	2.851996	68.15	0.000	188.7809	199.9606
cohor4Xaño06	367.841	3.140158	117.14	0.000	361.6863	373.9956
cohor5Xaño02	253.2959	2.752074	92.04	0.000	247.9019	258.6899
cohor5Xaño06	452.5233	3.10948	145.53	0.000	446.4288	458.6178
cohor6Xaño02	305.5361	2.800938	109.08	0.000	300.0463	311.0259
cohor6Xaño06	528.3737	3.13647	168.46	0.000	522.2263	534.5211
cohor7Xaño02	340.3865	2.886867	117.91	0.000	334.7283	346.0447
cohor7Xaño06	593.0545	3.283309	180.63	0.000	586.6193	599.4897
cohor8Xaño02	369.615	2.980523	124.01	0.000	363.7733	375.4568
cohor8Xaño06	643.3955	3.405899	188.91	0.000	636.72	650.071
cohor9Xaño02	404.9449	3.123556	129.64	0.000	398.8229	411.067
cohor9Xaño06	704.7743	3.674104	191.82	0.000	697.5732	711.9755
cohor10Xaño02	449.297	3.222535	139.42	0.000	442.981	455.6131
cohor10Xaño06	769.2403	3.855296	199.53	0.000	761.684	776.7966
cohor11Xaño02	502.6749	3.293174	152.64	0.000	496.2204	509.1295
cohor11Xaño06	850.1512	4.108037	206.95	0.000	842.0995	858.2028
cohor12Xaño02	538.0395	3.690384	145.79	0.000	530.8064	545.2725
cohor12Xaño06	890.9823	4.843418	183.96	0.000	881.4893	900.4752
_cons	18.49298	7.985392	2.32	0.021	2.841824	34.14414

## MUJERES

Number of obs = 105361  
 F(35, 105325) = 304.27  
 Prob > F = 0.0000  
 R-squared = 0.0918  
 Adj R-squared = 0.0915  
 Root MSE = 3.7655

Y = S	COEF.	STD. ERR.	t	P >  t	[95% CONF. INTERVAL]	
cohorte2	.6511572	.2599434	2.50	0.012	.1416717	1.160643
cohorte3	1.445609	.2494134	5.80	0.000	.9567623	1.934456
cohorte4	2.346819	.2472759	9.49	0.000	1.862162	2.831477
cohorte5	2.722275	.2474969	11.00	0.000	2.237185	3.207366
cohorte6	2.401856	.2490565	9.64	0.000	1.913709	2.890003

Anexo 2 (continuación)

### MODELO DE MOFFITT: PRIMERA ETAPA DE LA REGRESIÓN POR VARIABLES INSTRUMENTALES

Y = S	COEF.	STD. ERR.	t	P >  t	[95% CONF. INTERVAL]	
cohorte7	1.301289	.2492033	5.22	0.000	.8128537	1.789724
cohorte8	.5353468	.2499279	2.14	0.032	-.0454916	1.025202
cohorte9	.3078563	.2518153	1.22	0.222	-.1856982	.8014109
cohorte10	.1423358	.2550948	0.56	0.577	-.3576466	.6423182
cohorte11	-.1451086	.2588821	-0.56	0.575	-.652514	.3622968
cohorte12	-.5550068	.2644024	-2.10	0.036	-1.073232	-.0367816
cohor1Xaño02	1.928095	.2465462	7.82	0.000	1.444868	2.411322
cohor1Xaño06	2.906667	.2468725	11.77	0.000	2.422801	3.390534
cohor2Xaño02	2.277107	.1094094	20.81	0.000	2.062666	2.491548
cohor2Xaño06	2.584514	.1124242	22.99	0.000	2.364164	2.804864
cohor3Xaño02	1.572393	.0820399	19.17	0.000	1.411596	1.73319
cohor3Xaño06	1.399537	.0876328	15.97	0.000	1.227778	1.571296
cohor4Xaño02	.0501556	.0788332	0.64	0.525	-.1043563	.2046676
cohor4Xaño06	-.0970497	.0849402	-1.14	0.253	-.2635315	.069432
cohor5Xaño02	-.7881906	.0827793	-9.52	0.000	-.9504369	-.6259444
cohor5Xaño06	-1.290272	.088749	-14.54	0.000	-1.464219	-1.116325
cohor6Xaño02	-1.091187	.089745	-12.16	0.000	-1.267086	-.9152884
cohor6Xaño06	-1.4369	.0955055	-15.05	0.000	-1.62409	-1.24971
cohor7Xaño02	-.770445	.0924382	-8.33	0.000	-.9516226	-.5892674
cohor7Xaño06	-.9624807	.0988029	-9.74	0.000	-1.156133	-.7688284
cohor8Xaño02	-.4179737	.0958986	-4.36	0.000	-.6059336	-.2300138
cohor8Xaño06	-.3458786	.1044782	-3.31	0.001	-.5506545	-.1411027
cohor9Xaño02	-.4072232	.1063895	-3.83	0.000	-.6157452	-.1987012
cohor9Xaño06	-.2716844	.118877	-2.29	0.022	-.5046817	-.0386871
cohor10Xaño02	-.3041796	.1224572	-2.48	0.013	-.5441941	-.0641651
cohor10Xaño06	-.506209	.1388092	-3.65	0.000	-.7782731	-.2341449
cohor11Xaño02	-.5686028	.1398765	-4.07	0.000	-.8427589	-.2944467
cohor11Xaño06	-.3541704	.1616418	-2.19	0.028	-.6709862	-.0373545
cohor12Xaño02	-.2928679	.1664433	-1.76	0.078	-.6190945	.0333588
cohor12Xaño06	-.0266796	.1966927	-0.14	0.892	-.4121947	.3588355
_cons	7.74898	.2405668	32.21	0.000	7.277472	8.220487

Anexo 2 (continuación)

## MODELO DE MOFFITT: PRIMERA ETAPA DE LA REGRESIÓN POR VARIABLES INSTRUMENTALES

Number of obs = 105361  
 F(35, 105325) = 20290.87  
 Prob > F = 0.0000  
 R-squared = 0.8708  
 Adj R-squared = 0.8708  
 Root MSE = 3.8510

Y = EXPER	COEF.	STD. ERR.	t	P >  t	[95% CONF. INTERVAL]	
cohorte2	2.126961	.2658493	8.00	0.000	1.6059	2.648022
cohorte3	4.194933	.2550801	16.45	0.000	3.694979	4.694886
cohorte4	6.175383	.2528941	24.42	0.000	5.679714	6.671052
cohorte5	8.737368	.25312	34.52	0.000	8.241256	9.233479
cohorte6	12.05404	.2547151	47.32	0.000	11.5548	12.55328
cohorte7	16.16103	.2548652	63.41	0.000	15.66149	16.66056
cohorte8	19.87017	.2556062	77.74	0.000	19.36919	20.37116
cohorte9	23.08038	.2575365	89.62	0.000	22.57561	23.58515
cohorte10	26.17912	.2608905	100.35	0.000	25.66778	26.69046
cohorte11	29.55344	.2647639	111.62	0.000	29.03451	30.07238
cohorte12	32.94736	.2704097	121.84	0.000	32.41737	33.47736
cohor1Xaño02	4.664434	.2521477	18.50	0.000	4.170228	5.15864
cohor1Xaño06	7.678623	.2524814	30.41	0.000	7.183762	8.173483
cohor2Xaño02	4.462345	.1118951	39.88	0.000	4.243032	4.681658
cohor2Xaño06	8.190289	.1149785	71.23	0.000	7.964933	8.415646
cohor3Xaño02	5.243446	.0839038	62.49	0.000	5.078995	5.407896
cohor3Xaño06	9.569646	.0896238	106.78	0.000	9.393985	9.745308
cohor4Xaño02	6.856809	.0806243	85.05	0.000	6.698786	7.014831
cohor4Xaño06	11.15344	.0868701	128.39	0.000	10.98317	11.3237
cohor5Xaño02	7.790156	.08466	92.02	0.000	7.624223	7.956088
cohor5Xaño06	12.45441	.0907654	137.22	0.000	12.27651	12.63231
cohor6Xaño02	8.064972	.091784	87.87	0.000	7.885077	8.244868
cohor6Xaño06	12.61226	.0976754	129.12	0.000	12.42082	12.80371
cohor7Xaño02	7.702971	.0945384	81.48	0.000	7.517677	7.888265
cohor7Xaño06	12.14786	.1010477	120.22	0.000	11.94981	12.34591
cohor8Xaño02	7.342692	.0980774	74.87	0.000	7.150462	7.534923
cohor8Xaño06	11.57944	.106852	108.37	0.000	11.37001	11.78886
cohor9Xaño02	7.36556	.1088067	67.69	0.000	7.152301	7.57882
cohor9Xaño06	11.53273	.1215779	94.86	0.000	11.29444	11.77102
cohor10Xaño02	7.27827	.1252395	58.11	0.000	7.032803	7.523738

Anexo 2 (continuación)

### MODELO DE MOFFITT: PRIMERA ETAPA DE LA REGRESIÓN POR VARIABLES INSTRUMENTALES

Y = EXPER	COEF.	STD. ERR.	t	P >  t	[95% CONF. INTERVAL]	
cohor10Xaño06	11.81246	.1419629	83.21	0.000	11.53421	12.0907
cohor11Xaño02	7.452084	.1430545	52.09	0.000	7.171699	7.732469
cohor11Xaño06	11.53733	.1653143	69.79	0.000	11.21332	11.86134
cohor12Xaño02	7.13844	.1702249	41.94	0.000	6.804801	7.472078
cohor12Xaño06	11.17954	.2011616	55.57	0.000	10.78527	11.57382
_cons	3.702041	.2460324	15.05	0.000	3.219821	4.184261

Number of obs = 105361  
 F(35, 105325) = 24568.58  
 Prob > F = 0.0000  
 R-squared = 0.8909  
 Adj R-squared = 0.8908  
 Root MSE = 168.4025

Y = EXPER2	COEF.	STD. ERR.	t	P >  t	[95% CONF. INTERVAL]	
cohorte2	21.79056	11.62542	1.87	0.061	-.9951067	44.57622
cohorte3	52.98645	11.15449	4.75	0.000	31.12381	74.8491
cohorte4	92.3353	11.0589	8.35	0.000	70.66001	114.0106
cohorte5	151.5403	11.06878	13.69	0.000	129.8456	173.235
cohorte6	245.6776	11.13853	22.06	0.000	223.8462	267.509
cohorte7	392.3074	11.14509	35.20	0.000	370.4632	414.1517
cohorte8	552.7266	11.1775	49.45	0.000	530.8189	574.6344
cohorte9	713.7829	11.26191	63.38	0.000	691.7097	735.8561
cohorte10	889.0066	11.40858	77.92	0.000	866.6459	911.3673
cohorte11	1101.397	11.57796	95.13	0.000	1078.705	1124.09
cohorte12	1339.145	11.82484	113.25	0.000	1315.968	1362.321
cohor1Xaño02	65.7258	11.02626	5.96	0.000	44.11448	87.33712
cohor1Xaño06	131.9002	11.04085	11.95	0.000	110.2602	153.5401
cohor2Xaño02	83.66123	4.893103	17.10	0.000	74.07082	93.25165
cohor2Xaño06	178.2702	5.027937	35.46	0.000	168.4155	188.1249
cohor3Xaño02	120.6953	3.669061	32.90	0.000	113.504	127.8866
cohor3Xaño06	257.0323	3.919192	65.58	0.000	249.3507	264.7138
cohor4Xaño02	188.1105	3.525647	53.35	0.000	181.2003	195.0208
cohor4Xaño06	354.0157	3.798773	93.19	0.000	346.5701	361.4612
cohor5Xaño02	257.6571	3.702128	69.60	0.000	250.401	264.9133
cohor5Xaño06	471.1018	3.969111	118.69	0.000	463.3224	478.8812
cohor6Xaño02	320.9696	4.013656	79.97	0.000	313.1028	328.8363
cohor6Xaño06	560.8395	4.271284	131.30	0.000	552.4679	569.2112

Anexo 2 (continuación)

### MODELO DE MOFFITT: PRIMERA ETAPA DE LA REGRESIÓN POR VARIABLES INSTRUMENTALES

Y = EXPER2	COEF.	STD. ERR.	t	P >  t	[95% CONF. INTERVAL]	
cohor7Xaño02	364.6787	4.134104	88.21	0.000	356.5759	372.7815
cohor7Xaño06	631.3971	4.418752	142.89	0.000	622.7364	640.0578
cohor8Xaño02	399.6117	4.288862	93.17	0.000	391.2056	408.0179
cohor8Xaño06	681.7492	4.672569	145.90	0.000	672.591	690.9073
cohor9Xaño02	447.4834	4.758047	94.05	0.000	438.1577	456.8091
cohor9Xaño06	753.7368	5.316523	141.77	0.000	743.3165	764.1571
cohor10Xaño02	487.5127	5.476642	89.02	0.000	476.7786	498.2469
cohor10Xaño06	846.6753	6.207948	136.39	0.000	834.5078	858.8428
cohor11Xaño02	550.0452	6.255684	87.93	0.000	537.7842	562.3063
cohor11Xaño06	903.3769	7.229091	124.96	0.000	889.2079	917.5458
cohor12Xaño02	572.3876	7.443827	76.89	0.000	557.7978	586.9774
cohor12Xaño06	945.3507	8.796668	107.47	0.000	928.1093	962.592
_cons	16.22449	10.75884	1.51	0.132	-4.862696	37.31168

Fuente: Cálculos propios a partir de la EES.

## Anexo 3

## ESTIMACIÓN DE LAS ECUACIONES SALARIALES Y = LN (SALARIO)

	LICENCIADOS			LICENCIADOS + DIPLOMADOS		
	1995	2002	2006	1995	2002	2006
Exper	0,0349 (84,50)	0,0196 (52,02)	0,0155 (35,84)	0,0355 (86,05)	0,0220 (58,33)	0,0170 (38,29)
Exper2	-0,0005 (-61,54)	-0,0002 (-29,08)	-0,0002 (-21,03)	-0,0005 (-60,94)	-0,0002 (-32,34)	-0,0002 (-21,90)
Mujer	-0,1782 (-60,72)	-0,2087 (-81,52)	-0,1939 (-68,61)	-0,1639 (-55,79)	-0,2016 (-79,33)	-0,1893 (-66,43)
Extractivos	1,6932 (147,4)	1,9517 (186,6)	2,0204 (172,2)	1,6459 (146,0)	1,8903 (182,0)	1,9745 (169,8)
Alimentos y bebidas	1,7145 (254,6)	1,9089 (297,1)	1,9662 (263,6)	1,6690 (250,3)	1,8440 (290,5)	1,9145 (255,2)
Textil y confección	1,4982 (212,9)	1,7347 (256,1)	1,8225 (202,1)	1,4613 (210,8)	1,6761 (252,0)	1,7728 (197,9)
Cuero y calzado	1,5122 (192,2)	1,7273 (222,5)	1,7477 (156,7)	1,4850 (190,4)	1,6771 (218,2)	1,7066 (152,2)
Madera y corcho	1,4542 (195,0)	1,7436 (259,1)	1,8222 (222,3)	1,4260 (193,2)	1,6931 (256,6)	1,7801 (216,4)
Papel	1,7957 (225,5)	2,0667 (257,7)	2,1044 (219,4)	1,7487 (222,1)	1,9976 (252,1)	2,0528 (210,5)
Industria química	1,9984 (252,1)	2,2259 (258,6)	2,2761 (208,5)	1,9391 (246,4)	2,1471 (250,6)	2,1955 (197,3)
Caucho y materia	1,6942 (205,3)	1,9775 (259,0)	2,0520 (216,2)	1,6547 (204,0)	1,9157 (255,6)	2,0020 (209,4)
Minerales no metálico	1,7232 (241,6)	1,9479 (293,7)	2,0410 (266,4)	1,6781 (238,7)	1,8868 (287,5)	1,9887 (260,5)
Metalurgia y productos metálicos	1,7338 (257,5)	1,9941 (344,9)	2,0413 (306,6)	1,6935 (253,1)	1,9354 (337,9)	1,9931 (295,2)
Construcción de maquinaria	1,7884 (229,8)	2,0762 (302,9)	2,1076 (228,1)	1,7378 (226,0)	2,0049 (295,6)	2,0471 (218,3)
Material y equipo eléctrico	1,8080 (235,2)	2,0785 (297,6)	2,1327 (224,3)	1,7357 (231,5)	1,9989 (293,1)	2,0600 (216,4)
Fabricación de material transporte	1,8044 (251,3)	2,1165 (320,0)	2,2015 (249,1)	1,7576 (248,4)	2,0511 (314,4)	2,1451 (241,7)
Manufactureras diversas	1,5426 (219,3)	1,7666 (279,0)	1,8509 (260,5)	1,5101 (217,2)	1,7134 (274,3)	1,8094 (253,0)
Energía eléctrica, gas y agua	2,0466 (195,0)	2,3094 (215,9)	2,4346 (173,4)	1,9724 (187,0)	2,2041 (199,6)	2,3456 (164,5)

Anexo 3 (continuación)

### ESTIMACIÓN DE LAS ECUACIONES SALARIALES Y = LN (SALARIO)

	LICENCIADOS			LICENCIADOS + DIPLOMADOS		
	1995	2002	2006	1995	2002	2006
Construcción	1,6219 (253,3)	1,8601 (352,5)	1,9467 (335,9)	1,5583 (249,8)	1,7847 (347,2)	1,8826 (320,6)
Comercio	1,6706 (285,7)	1,9289 (370,8)	1,9843 (347,3)	1,6301 (281,6)	1,8655 (363,8)	1,9401 (331,9)
Hostelería	1,5739 (240,4)	1,7974 (323,5)	1,8840 (293,2)	1,5322 (234,7)	1,7395 (314,5)	1,8375 (278,8)
Transporte	1,8385 (257,9)	2,0314 (315,6)	2,0721 (284,7)	1,7844 (251,3)	1,9556 (312,4)	2,0113 (272,1)
Intermediación financiera	2,2424 (333,1)	2,5707 (375,1)	2,6734 (320,0)	2,2025 (314,5)	2,5064 (340,8)	2,6310 (279,5)
Actividades inmobiliarias	1,6849 (222,0)	1,7989 (331,8)	1,8946 (314,6)	1,6113 (213,8)	1,7192 (319,8)	1,8308 (295,2)
N	131.330	137.334	100.981	124.501	128.956	93.990
R2 ajustado	0,964	0,968	0,970	0,967	0,971	0,972

Fuente: Cálculos propios a partir de la EES.

() t-estadísticos robustos a la heteroscedasticidad.



## Últimos números publicados

- N.º 30. LA INDUSTRIA DE ALTA TECNOLOGÍA EN ESPAÑA: FACTORES DE LOCALIZACIÓN Y DINÁMICA ESPACIAL (*Serie TESIS*),  
por Miguel Giner Pérez.
- N.º 31. CONVERGENCIA EN RENTA PER CÁPITA ENTRE LAS COMUNIDADES AUTÓNOMAS ESPAÑOLAS (1955-2004): UNA APLICACIÓN BASADA EN MÉTODOS DE PANEL DINÁMICO (*Serie TESIS*),  
por Fernando Martín Mayoral.
- N.º 32. EL DESDOBLAMIENTO DE ACCIONES EN EL MERCADO ESPAÑOL: FACTORES DETERMINANTES Y EFECTOS (*Serie TESIS*),  
por María Eugenia Ruiz Molina.
- N.º 33. EL TRABAJO DOMÉSTICO CUENTA: LAS CUENTAS DE LOS HOGARES EN ESPAÑA 1996 Y 2003 (*Serie ECONOMÍA Y SOCIEDAD*),  
por María Luisa Moltó y Ezequiel Uriel.
- N.º 34. GESTIÓN DEL MEDIO NATURAL EN LA PENÍNSULA IBÉRICA: ECONOMÍA Y POLÍTICAS PÚBLICAS (*Serie ECONOMÍA Y SOCIEDAD*),  
por Pablo Campos Palacín y José-María Casado Raigón.
- N.º 35. PATRIMONIO INMOBILIARIO Y BALANCE NACIONAL DE LA ECONOMÍA ESPAÑOLA (1995-2007) (*Serie ECONOMÍA Y SOCIEDAD*),  
por José Manuel Naredo, Óscar Carpintero y Carmen Marcos.
- N.º 36. EN TORNO A LA FAMILIA ESPAÑOLA: ANÁLISIS Y REFLEXIONES DESDE PERSPECTIVAS SOCIOLOGICAS Y ECONÓMICAS (*Serie ECONOMÍA Y SOCIEDAD*),  
por Elisa Chuliá y José Félix Sanz (coordinadores).
- N.º 37. PROBLEMÁTICA DE LA DEPENDENCIA EN ESPAÑA: ASPECTOS DEMOGRÁFICOS Y DEL MERCADO DE TRABAJO (*Serie ECONOMÍA Y SOCIEDAD*),  
por Lorenzo Serrano y Ángel Soler.
- N.º 38. EDUCACIÓN Y FAMILIA. LOS PADRES ANTE LA EDUCACIÓN GENERAL DE SUS HIJOS EN ESPAÑA (*Serie ECONOMÍA Y SOCIEDAD*),  
por Víctor Pérez-Díaz, Juan Carlos Rodríguez y Juan Jesús Fernández.
- N.º 39. COMPETITIVIDAD Y DESLOCALIZACIÓN EN LA INDUSTRIA ESPAÑOLA (*Serie ECONOMÍA Y SOCIEDAD*),  
por Diego Rodríguez, Jaime Turrión y Francisco J. Velázquez.
- N.º 40. DOS ENSAYOS SOBRE FINANCIACIÓN AUTONÓMICA (*Serie ECONOMÍA Y SOCIEDAD*),  
por Carlos Monasterio Escudero e Ignacio Zubiri Oria.
- N.º 41. EFICIENCIA Y CONCENTRACIÓN DEL SISTEMA BANCARIO ESPAÑOL (*Serie ANÁLISIS*),  
por Fernando Maravall, Silviu Glavan y Analistas Financieros Internacionales.
- N.º 42. ANÁLISIS DE REFORMAS DEL IMPUESTO SOBRE LA RENTA PERSONAL A PARTIR DE MICRODATOS TRIBUTARIOS (*Serie ANÁLISIS*),  
por José Félix Sanz, Juan Manuel Castañer Carrasco y Desiderio Romero Jordán.
- N.º 43. COMPORTAMIENTO ESTRATÉGICO DE LA BANCA AL POR MENOR EN ESPAÑA: FUSIONES Y ESPECIALIZACIÓN GEOGRÁFICA (*Serie TESIS*),  
por Cristina Bernad Morcate
- N.º 44. LA VERTIENTE CUALITATIVA DE LA MATERIALIDAD EN AUDITORÍA: MARCO TEÓRICO Y ESTUDIO EMPÍRICO PARA EL CASO ESPAÑOL (*Serie TESIS*),  
por Javier Montoya del Corte
- N.º 45. LA DECISIÓN DE INTERNACIONALIZACIÓN DE LAS EMPRESAS: UN MODELO TEÓRICO CON INVERSIÓN HORIZONTAL Y VERTICAL (*Serie TESIS*),  
por Jaime Turrión Sánchez

- N.º 46. **FINANCIACIÓN DE LA ENSEÑANZA OBLIGATORIA: LOS BONOS ESCOLARES EN LA TEORÍA Y EN LA PRÁCTICA** (*Serie ECONOMÍA Y SOCIEDAD*),  
por Javier Díaz Malledo (coordinador), Clive R. Belfield, Henry M. Levin, Alejandra Mizala, Anders Böhlmark, Mikael Lindahl, Rafael Granell Pérez y María Jesús San Segundo
- N.º 47. **SERVICIOS Y REGIONES EN ESPAÑA** (*Serie ECONOMÍA Y SOCIEDAD*),  
por Juan R. Cuadrado Roura y Andrés Maroto Sánchez
- N.º 48. **LAS EMPRESAS DEL SECTOR DE LA CONSTRUCCIÓN E INMOBILIARIO EN ESPAÑA: DEL BOOM A LA RECESIÓN ECONÓMICA** (*Serie ECONOMÍA Y SOCIEDAD*),  
por Belén Gill de Albornoz (Dir.), Juan Fernández de Guevara, Begoña Giner y Luis Martínez
- N.º 49. **INSTRUMENTOS PARA MEJORAR LA EQUIDAD, TRANSPARENCIA Y SOSTENIBILIDAD DE LOS SISTEMAS DE PENSIONES DE REPARTO** (*Serie TESIS*),  
por M.ª del Carmen Boado-Penas
- N.º 50. **EL IMPUESTO DE FLUJOS DE CAJA EMPRESARIAL: UNA ALTERNATIVA AL IMPUESTO SOBRE LA RENTA DE SOCIEDADES** (*Serie TESIS*),  
por Lourdes Jerez Barroso
- N.º 51. **LA SUBCONTRATACIÓN DE SERVICIOS DE I+D: EVIDENCIA DE EMPRESAS EUROPEAS Y DE EE.UU.** (*Serie TESIS*),  
por Andrea Martínez Noya
- N.º 52. **IMPOSICIÓN EFECTIVA SOBRE LAS RENTAS DEL CAPITAL CORPORATIVO: MEDICIÓN E INTERPRETACIÓN. EL IMPUESTO DE SOCIEDADES EN ESPAÑA Y EN LOS PAÍSES DE LA UNIÓN EUROPEA EN EL CAMBIO DE MILENIO** (*Serie ANÁLISIS*),  
por José Félix Sanz Sanz, Desiderio Romero Jordán y Begoña Barruso Castillo
- N.º 53. **¿ES RENTABLE EDUCARSE? MARCO CONCEPTUAL Y PRINCIPALES EXPERIENCIAS EN LOS CONTEXTOS ESPAÑOL, EUROPEO Y EN PAÍSES EMERGENTES** (*Serie ECONOMÍA Y SOCIEDAD*),  
por José Luis Raymond (coordinador)

# ESTUDIOS DE LA FUNDACIÓN

SERIE ECONOMÍA Y SOCIEDAD

*Pedidos e información:*

FUNDACIÓN DE LAS CAJAS DE AHORROS

Caballero de Gracia, 28  
28013 Madrid

Teléfono: 91 596 54 81

Fax: 91 596 57 96

[suscrip@funcas.es](mailto:suscrip@funcas.es)

[www.funcas.es](http://www.funcas.es)

P.V.P.: 13 € (IVA incluido)

ISBN 978-84-89116-70-2



9 788489 116702