



# Los rendimientos de la educación desde un enfoque agregado: una evaluación con datos regionales

Juan Antonio Duro Moreno  
Universidad de Barcelona

BIBLID [0213-7525 (1997): 49: 15-33]

PALABRAS CLAVE: Tasas de rendimiento de la educación, nivel de desarrollo.

## RESUMEN:

El trabajo examina empíricamente las tasas de rendimiento de la educación desde un enfoque agregado, en línea con la propuesta de Ram (1996). Los beneficios a nivel práctico derivados de este enfoque son relevantes. La principal novedad de este trabajo radica en la utilización de la región española como unidad básica de análisis. Adicionalmente, se abordan algunos aspectos de interés relativos al patrón global de las tasas de rendimiento. El trabajo empírico se ha visto facilitado por la publicación de nuevos datos territorializados sobre escolarización de la población.

## ABSTRACT:

The paper empirically analyzes the returns to investment in education in a global context, following the work by Ram (1996). The practical advantages derived from this work are relevant. Originally, the analytical basic unit considered are the Spanish regions. Additionally, we analyze some aspects of interest on the global pattern of the returns to education. The empirical work has been facilitated by the publication of new territorialized population schooling data.

---

## 1. LÓGICA DEL ESTUDIO

---

La estimación de tasas de rendimiento en educación se ha convertido en uno de los campos de investigación más fructíferos de los últimos treinta años<sup>1</sup>. Dos métodos han sido generalmente aplicados en orden a estimar la rentabilidad de las inversiones educativas: el "método elaborado" y el "método de las funciones generadoras de ingresos". El "método elaborado" consiste en un

1. Para una revisión de la literatura el lector interesado podría consultar, entre otros, el artículo recopilatorio de Schultz (1988)

análisis coste-beneficio, donde la tasa de retorno es calculada a través de la confrontación de la corriente de costes asociados a la educación (básicamente, los costes de oportunidad) con la de beneficios en un punto temporal determinado. Este procedimiento permite computar no sólo las tasas privadas de rendimiento sino también las tasas sociales, dependiendo del tipo de costes y beneficios considerados en el cálculo<sup>2</sup>. Tuvo una amplia difusión durante la década de los sesenta, gracias a las aportaciones de Schultz (1961) y Becker (1964). Tras la publicación del conocido trabajo de Mincer (1974), sin embargo, la literatura se ha inclinado por la utilización de las funciones generadoras de ingresos, menos exigentes en términos de datos, y, por tanto, más fáciles de emplear. Su aplicación se encuentra asociada al ajuste de una regresión semi-logarítmica del logaritmo de los ingresos individuales con los años de escolarización y la experiencia laboral, introducida ésta última variable tanto de forma lineal como cuadrática. Esta técnica sólo permite estimar, de manera directa, la tasa promedio privada de rendimiento<sup>3</sup>.

Para el caso específico español, los trabajos empíricos más actualizados se han centrado en la aplicación de la segunda de las metodologías. En este sentido, es conveniente destacar las contribuciones de García Calvo (1988), Lassibille (1988) o Alba y San Segundo (1995), entre otros.

Recientemente, Ram (1996) ha explotado la metodología de Mincer con el objetivo de producir estimaciones sobre tasas agregadas de rentabilidad educativa. Específicamente, se trataría de utilizar las especificaciones convencionales micro pero empleando, como unidad básicas de estudio, los países, en lugar de los individuos. Este procedimiento genera ventajas operativas apreciables en relación al empleo de datos individualizados. En primer lugar, la utilización de información de base agregada disminuye la necesidad de incluir en las ecuaciones variables que son de difícil cuantificación y disponibilidad individual, como la experiencia laboral o la habilidad. Segundo, la utilización de datos agregados permite capturar, ni que sea burdamente, las posibles externalidades asociadas a la educación, factor no considerado en las usuales ecuaciones de Mincer<sup>4</sup>. Finalmente, al considerar como variable

2. Para el cálculo social es necesario aumentar los costes privados con los costes globales de los recursos, principalmente públicos, y en la corriente de beneficios deben incluirse, adicionalmente, las posibles externalidades asociadas a la educación.
3. En Psacharopoulos (1980) se puede encontrar una instructiva discusión entorno a ambas metodologías.
4. Estas externalidades tampoco son generalmente consideradas en las estimaciones basadas en el método elaborado. De hecho, el cálculo usual de las tasas sociales se realiza simplemente adicionando a los costes privados los gastos públicos en educación. Por esta razón, es común observar que las tasas sociales son inferiores a las tasas privadas. Ver, por ejemplo, las últimas estimaciones de Psacharopoulos (1994).

dependiente una medida de la productividad agregada, este enfoque permite analizar explícitamente el efecto de la educación sobre la productividad, en consistencia con las teorías del capital humano.

Este trabajo se enmarca dentro del enfoque agregado. Su principal innovación, respecto al pionero trabajo de Ram (1996), emerge del empleo de información infranacional. La consideración de datos regionales produce algunos beneficios dignos de mención. Los principales vienen derivados de su superior comparabilidad, y calidad, frente a los datos internacionales comúnmente utilizados<sup>5</sup>.

En consecuencia, el principal propósito de este trabajo es proporcionar evidencia empírica sobre la estimación de tasas agregadas de rendimiento en educación partiendo de datos regionales para España, que cubren el periodo 1964-1991. Adicionalmente, se evalúan algunas hipótesis frecuentemente examinadas en la literatura internacional entorno al patrón de las tasas de rendimiento.

---

## 2. ESPECIFICACIÓN DEL MODELO Y DATOS

---

Ram (1996) utiliza, en orden a generar estimaciones agregadas de las tasas de rendimiento, el modelo de escolarización propuesto por Mincer (1974). Aumentando éste con variables dummies por país la ecuación a estimar queda como sigue:

$$\ln(Y_{it}) = \alpha + \beta EM_{it} + \sum_{i=2}^n \delta_i D_i + \mu_{it} \quad (1)$$

donde  $Y_{it}$  sería la renta por empleado en el país  $i$  en el año  $t$ ,  $EM_{it}$  representaría los años de escolarización promedio por empleado en el país  $i$  en el año  $t$ ,  $D_i$  sería una variable dummy que tomaría el valor 1 para el país  $i$  y 0 para el resto y  $u_{it}$  denotaría la clásica perturbación aleatoria convencionalmente incluida en los modelos de regresión. Las variables ficticias se incluyen en la ecuación para capturar la influencia de otras variables sobre las rentas por ocupado<sup>6</sup>. El modelo puede ser adaptado fácilmente en términos regionales.

5. El propio Ram (1996) expone algunas de las limitaciones características de los datos por países.
6. Entre otros aspectos, las variables ficticias pueden reflejar las diferencias en los costes de la educación.

El parámetro  $\alpha$  se interpretaría, siguiendo la lógica minceriana, como el valor de los ingresos medios que obtendría un país con cero años de escolarización<sup>7</sup> y el parámetro  $\beta$  denotaría una especie de tasa social media de rendimiento en educación<sup>8</sup>, por cuanto incluiría el efecto de las externalidades vinculadas a la educación. Obsérvese que la interpretación de  $\beta$  es conceptualmente distinta a la indicada en los modelos tipo-Mincer convencionales, donde la  $\beta$  se identifica como la tasa privada promedio de rendimiento educativo.

No obstante, la estimación de tasas de rendimiento agregadas, ya sean regionales o nacionales, en base a un modelo como (1) puede ser insatisfactoria debido a la sospecha de la presencia de errores de especificación significativos, que pueden sesgar las estimaciones. Específicamente, la principal fuente de sesgo puede ser atribuible a un problema de omisión de variables relevantes. Como es bien sabido, aparecerán distorsiones en las estimaciones cuando las variables omitidas estén estadísticamente correlacionadas con las variables incluidas. En este sentido, una de las primeras variables candidatas a incluir en la ecuación sería el stock de capital por trabajador<sup>9</sup>. En nuestros datos regionales, EM y el stock de capital por trabajador exhiben un coeficiente de correlación parcial del 0.82. Esta elevada colinearidad entre ambas variables, de hecho, produce una sobreestimación de  $\beta$ , cuando es calculada siguiendo la especificación (1)<sup>10</sup>. Este fenómeno unido a la teóricamente plausible relación independiente entre renta y stock de capital aconsejan la utilización de un modelo más amplio que (1). En consecuencia, se postula el siguiente modelo modificado<sup>11</sup>:

$$\ln(Y_{it}) = \alpha + \beta EM_{it} + \gamma \ln(k_{it}) + \sum_{i=2}^n \delta_i D_i + \mu_{it} \quad (2)$$

donde  $k_{it}$  sería el stock de capital por trabajador en la región  $i$ <sup>12</sup>.

7. Nótese que al introducir las dummies en el modelo se está postulando explícitamente que este parámetro puede variar entre países
8. Esta tasa social derivada del empleo de datos agregados con el modelo de Mincer no es conceptualmente equivalente a la tasa estimada en base al método elaborado. La razón básica de esta diferenciación se atribuye a que la primera supone que los costes de oportunidad son el único componente de coste asociado a las inversiones educativas.
9. Esta es una variable frecuentemente considerada en los modelos de funciones de producción agregada. Ver, Solow para una racionalidad teórica y Mankiw, Romer y Weil (1992) para una aplicación empírica.
10. Estos cálculos fueron realizados por el autor, estando disponibles para cualquier petición
11. Ram (1996) comprueba la sensibilidad de sus estimaciones comparando los resultados desprendidos con y sin incluir en los ajustes el stock de capital. Entre otras cosas, no encuentra ninguna discrepancia significativa. Es oportuno mencionar, no obstante, que este análisis de robustez es abordado con una muestra muy corta, de tan sólo 22 observaciones. Nuestro análisis comparativo parece más eficiente, por cuanto se ha utilizado una muestra formada por 238 observaciones.

El empleo de un modelo como (2) para estimar tasas de rendimiento arroja algunas ventajas significativas frente a los modelos mincerianos convencionales. Primero, es bien sabido que un problema común de los estudios micro proviene de la omisión de una variable habilidad, omisión que puede causar un error en las estimaciones<sup>13</sup>. Esta omisión es generalmente atribuible a la difícil cuantificación, y disponibilidad, de información estadística sobre esta variable en las fuentes habituales. Este problema puede ser, sin embargo, menos relevante en un ámbito agregado ya que es menos probable que las diferencias económicas entre regiones se vean afectadas, de forma sistemática, por la habilidad innata de los ocupados. Segundo, las especificaciones individuales usualmente incluyen una variable "experiencia laboral", de forma lineal y cuadrática, como determinante de los ingresos individuales. En la amplia mayoría de los casos, al no disponerse de información cuantitativa, esta variable ha tenido que aproximarse<sup>14</sup>. Obviamente, la aproximación produce una imprecisión de difícil cuantificación. En nuestro caso, la introducción de esta variable parece, en consonancia con la argumentación anterior, menos necesaria. Finalmente, obsérvese que el coeficiente  $\beta$  refleja, en realidad, el efecto de la educación sobre la productividad, que es precisamente el efecto abordado teóricamente por la literatura de capital humano. Así, resulta irrelevante la discusión acerca de la supuesta asociación entre niveles de ingresos y productividad, muy usual en los modelos micro.

Por otra parte, a partir del modelo agregado básico, ecuación (2), y nuestros datos puede ser útil analizar algunos aspectos interesantes vinculados al patrón de las tasas de rendimiento. Entre ellos, puede investigarse la validez empírica de la hipótesis de los rendimientos decrecientes en las inversiones educativas y, asimismo, evaluar la posible relación existente entre tasas de rendimiento y nivel de desarrollo.

En cuanto a los datos regionales utilizados es conveniente señalar, ante todo, que su grado comparabilidad parece superior al existente a un nivel internacional. Los datos regionalizados sobre educación de la población ocupada proceden de una única fuente: la reconstrucción temporal realizada recientemente por Mas, Pérez, Uriel y Serrano (1995)<sup>15</sup> y los de renta se han

---

12. En la sección IV del trabajo se indica que el modelo (2) no padece ningún problema adicional asociado a la omisión de variables relevantes.

13. Griliches (1977) fue uno de los autores principales encargados en enfatizar este problema

14. Mincer (1974) sugiere un posible método para aproximar la experiencia laboral. Concretamente, se trataría de restar a la edad del individuo un factor de corrección que varía en función del nivel de enseñanza alcanzado.

15. En adelante, los vamos a denominar MPUS.

tomado de las ya clásicas estimaciones realizadas por el BBV. En este sentido, se ha considerado el PIB como indicador de renta.<sup>16</sup> Los datos sobre stocks de capital provienen de la Fundación BBV. La conjugación de todas las fuentes ha permitido disponer de datos bianuales por regiones para el período 1964-1991.

La variable de educación media de la población (EM) ha sido construida siguiendo una metodología bastante estándar. Así, el nivel de escolarización medio de la población ocupada por región se calcula a partir de la siguiente expresión:

$$EM_i = \sum_j L_{ij} E_j \quad (3)$$

donde  $L_{ij}$  es la proporción de la población ocupada que posee el nivel educativo  $j$  en la región  $i$ , y  $E_j$  es el número de años de escolarización medio atribuido al nivel educativo  $j$ . MPUS facilitan información estadística para las siguientes categorías educativas: analfabetos, sin estudios y primarios, medios, grado superior 1 ciclo y grado superior 2º ciclo. Los años atribuidos a cada nivel educativo han sido: analfabetos: 0 años, sin estudios y primarios: 3 años, medios: 10 años, grado superior 1 ciclo: 15 años y grado superior 2º ciclo: 17 años<sup>17</sup>.

La información disponible ha permitido contar con 238 observaciones, (17x14), número bastante aceptable desde el punto de vista de la eficiencia estimativa. El empleo de datos de panel permite, entre otros aspectos, introducir variables dummies regionales, y así recoger diferencias en rentas regionales no achacables ni al nivel de educación medio ni al stock de capital.

---

### 3. RESULTADOS EMPÍRICOS

---

A continuación se estima la ecuación (2) con toda la muestra disponible. Antes de todo, señalar que debido a los claros indicios de autocorrelación residual en las estimaciones se ha aplicado un método iterativo de corrección, en el que se postula que los errores siguen un proceso AR(1). Entre paréntesis

16. Otras variables podrían haber sido consideradas alternativamente, como es el caso de la renta regional. En cualquier caso, la utilización de una u otra medida no genera cambios significativos en los resultados, como ha comprobado el autor.
17. Una de las principales limitaciones de estos datos está asociada con la excesiva agregación de los grupos educativos, que disminuye la precisión en los cálculos de EM. Esto es particularmente relevante en el caso de los estudios medios. Para un repaso entorno a algunos de los principales problemas derivados del empleo de los datos MPUS puede consultarse Duro (1997).

figuran los valores del estadístico t, que contrasta la significatividad de los parámetros. Los asteriscos indican significatividad estadística de los coeficientes a un nivel del 1%. Las variables dummies y el término autorregresivo de primer orden no han sido reproducidos para salvar espacio.

$$\ln(Y) = -0,5581 + 0,0579 \text{ EM} + 0,6583 \ln(k) + \dots$$

(16,92)\*    (7,56)\*            (29,21)\*

$$\bar{R}^2 = 0,98, \text{ SEE} = 0,4576, \text{ D.W.} = 1,95$$

A la luz de los resultados se desprende que la tasa de rendimiento es positiva, significativa y próxima al 6%<sup>18</sup>. Aunque la comparabilidad entre éste resultado y los obtenidos en otros estudios no es estricta, debido al uso de diferentes metodologías, datos y fuentes, puede ser útil examinar la evidencia encontrada por otros autores para España. García Calvo (1988) estima los rendimientos de la educación por niveles educativos con datos referentes a cabezas de familia recogidos de la EPF-81. En lugar de emplear una variable continua como los años de escolarización promedio utiliza una serie de variables dummies vinculadas a la terminación de los principales ciclos educativos (primaria, secundaria y terciaria)<sup>19</sup>. A partir de los coeficientes de las variables dummies es posible derivar una tasa de rendimiento promedio, en línea con la propuesta de Psacharopoulos (1980). En particular, la tasa promedio estimada a partir de los datos de García Calvo sería del 6,3%, cifra muy cercana a la que se desprende del análisis agregado regional. Lassibille (1988), por su parte, ajusta una función de ingresos similar para Andalucía con datos por cabezas de familia, asalariados y empresarios, procedentes también de la EPF-81. Entre sus principales hallazgos encuentra una tasa de rendimiento privado para esta región del 6,8% promedio, en el caso de los asalariados, y del 6,5%, para los empresarios. En otro trabajo, Alba y San Segundo (1995) aplican idéntica metodología a datos educativos cualitativamente superiores, procedentes de la EPA de 1990 (segundo trimestre)<sup>20</sup>. Según sus resultados, la

18. Puesto que el trabajo se centra en el análisis de las tasas agregadas de rendimiento, la discusión sobre los parámetros asociados a otras variables se omite.

19. Este es el denominado por Psacharopoulos (1994) "método extendido", puesto que permite calcular tasas de rendimiento por niveles educativos.

20. Entre otras razones, estos datos son de una calidad mejor porque suministran información para un mayor número de niveles educativos y no se restringe la información a los cabezas de familia.

tasa promedio privada sería del 6,6%. A diferencia de los anteriores, Psacharopoulos (1994) ofrece estimaciones de tasas de rendimiento siguiendo el método elaborado. Con datos de 1971 para España halla que la tasa social de rendimiento sería del 12,9%, un 18,9% si el cálculo fuera en términos privados. Nótese, en este último caso, que la antigüedad de los datos junto a un posible patrón decreciente de la tasa de rendimiento, hipótesis sostenida por el propio autor [Psacharopoulos (1989)], contribuirían a interpretar la elevada discrepancia. En consecuencia, la tasa agregada obtenida parece plausible y es consistente con las estimaciones producidas por otros autores, con otros métodos y datos

Llegados a este punto, una hipótesis que parece instructivo contrastar es si se observan o no rendimientos decrecientes en las inversiones educativas, al igual que puede ocurrir con otras formas de capital. Es decir, sería interesante investigar si se verifica que el producto marginal de las inversiones educativas es decreciente. Esta hipótesis puede ser fácilmente explorada, con el método agregado, a través de la especificación de una ecuación como la siguiente:

$$\ln(Y_{it}) = \alpha + \beta_1 EM_{it} + \beta_2 EM_{it}^2 + \gamma \ln(k_{it}) + \sum_{i=2}^n \delta_i D_i + \mu_{it} \quad (4)$$

El término cuadrático de la educación reflejaría el efecto de la expansión educativa en el margen. De tal forma que si se cumpliera la hipótesis de los rendimientos decrecientes en la educación el parámetro  $\beta_1$  habría de mostrar un signo positivo, y significativo, y el parámetro  $\beta_2$ , por su parte, habría de exhibir un signo negativo.

Los datos disponibles producen el siguiente resultado:

$$\ln(Y) = -0,9036 + 0,2352 EM - 0,0138 EM^2 + 0,5743 \ln(k) + \dots,$$

(8,14)\*      (4,31)\*      (3,28)\*      (16,69)\*

$$\bar{R}^2 = 0,98, \text{ SEE} = 0,4358, \text{ D.W.} = 1,96$$

Como se observa, los resultados revelan un claro apoyo a la hipótesis de los rendimientos decrecientes en la educación. Así, parece que las expansiones educativas, ceteris paribus, producen mayores rendimientos en las regiones con menores niveles educativos. Análogo patrón ha sido documentado por Ram (1996) utilizando un modelo parecido con datos internacionales. En concreto, la ecuación ajustada revela que la tasa de rendimiento es variable y

seguiría el patrón  $0,2352-0,0276*EM$ . Esto es, cada año adicional de escolarización media, todo lo demás constante, reduciría la tasa de rendimiento agregada en aproximadamente 2,8 puntos porcentuales.

Por otra parte, una atención especial ha sido dedicada en la literatura a la conexión entre tasas de rendimiento en educación y niveles de renta. Psacharopoulos (1994) encuentra, en términos generales, que las tasas de rendimiento micro (intranacionales) son mayores en los países cuyo grado de desarrollo es más reducido. Asimismo, Ram (1996) descubre que la tasa agregada de rendimiento en los países de desarrollo bajo y medio es sustancialmente mayor que la reflejada por los países desarrollados. Sin embargo, este autor encuentra, adicionalmente, que la rentabilidad educativa promedio es mayor en los países de desarrollo medio que en los países de desarrollo reducido. Tiene interés evaluar esta cuestión con nuestros datos. Para llevar a cabo el análisis es necesario segmentar la muestra, atendiendo a algún criterio del nivel de desarrollo. Dado el reducido número de regiones disponible se ha creído conveniente formar únicamente dos grupos, utilizando el PIB per cápita relativo como criterio de discriminación. En un primer grupo situaríamos a las regiones cuyo nivel de desarrollo, medido por el PIB per cápita relativo en promedio temporal, excede o está próximo a la media española y otro, estaría formado por las regiones cuyo nivel es claramente inferior<sup>21</sup>. Las estimaciones para cada uno de los subgrupos han sido corregidas por autocorrelación y las variables dummies han sido incluidas. Los resultados alcanzados han sido los siguientes:

*Regiones de desarrollo relativo medio y elevado*

$$\ln(Y) = -0,4317 + 0,06898 EM + 0,5893 \ln(k) + \dots$$

(11,63)\*    (7,79)\*    (20,26)\*

$$\bar{R}^2 = 0,98, \text{ SEE} = 0,2501, \text{ D.W.} = 1,98$$

$$\ln(Y) = -0,8219 + 0,2618 EM - 0,0145 EM^2 + 0,4923 \ln(k) + \dots$$

(5,71)\*    (3,82)\*    (2,84)\*    (10,99)\*

$$\bar{R}^2 = 0,98, \text{ SEE} = 0,2351, \text{ D.W.} = 1,99$$

21. En el primer grupo se incluye a Aragón, Asturias, Baleares, Cantabria, Cataluña, Comunidad Valenciana, Madrid, Navarra, País Vasco y La Rioja. El segundo grupo, por su parte, contiene a Andalucía, Canarias, Castilla-La Mancha, Castilla-León, Extremadura, Galicia y Murcia.

*Regiones de desarrollo relativo reducido*

$$\ln(Y) = -0,6329 + 0,05478 \text{ EM} + 0,6990 \ln(k) + \dots$$

(14,64)\*    (3,93)\*    (19,82)\*

$$\bar{R}^2 = 0,99, \text{ SEE} = 0,1632, \text{ D.W.} = 1,86$$

$$\ln(Y) = -1,2633 + 0,3724 \text{ EM} - 0,0276 \text{ EM}^2 + 0,5892 \ln(k) + \dots$$

(5,34)\*    (3,14)\*    (2,70)\*    (10,75)\*

$$\bar{R}^2 = 0,98, \text{ SEE} = 0,1502, \text{ D.W.} = 1,89$$

Con estos resultados se comprueba que la tasa de rendimiento agregada parece ser algo mayor, en promedio, en las regiones de desarrollo medio-elevado, 6,9% frente a 5,5%. Esta observación, sin ir más lejos, se asemeja a la encontrada por Ram (1996) al comparar el rendimiento agregado exhibido por los países de desarrollo medio con el reflejado por los países de desarrollo reducido. Obsérvese, por otra parte, que a pesar de la diferencia en los promedios, ambos grupos coinciden en exhibir el patrón de los rendimientos decrecientes.

Para abundar un poco más en la observación entorno a la discrepancia de rentabilidad educativa según el nivel de desarrollo, se han reestimado las tasas de rendimiento por región utilizando las series temporales. Es conveniente reseñar que éstas únicamente se componen de 14 observaciones, por lo que el análisis ha de ser considerado como una mera ilustración empírica<sup>22</sup>. El Cuadro 1 ofrece los resultados que se desprenden al ajustar la ecuación (2), lógicamente sin dummies, a cada una de las regiones. Entre otros aspectos, se constata que regiones como Baleares, Navarra o Cataluña, de elevado desarrollo económico comparativo, coinciden en reflejar tasas de rendimiento

22. Parece apropiado señalar que las ecuaciones región por región podrían haber sido estimadas conjuntamente como un sistema de ecuaciones aparentemente no relacionadas (o SURE). Esta metodología es útil en los casos en los que los términos de error entre ecuaciones están correlacionados (piénsese, por ejemplo, en el efecto de los shocks temporales sobre las rentas regionales). Esta posibilidad llevó a intentar estimar las ecuaciones como un sistema, siguiendo el método de Zellner. Sin embargo, esto no fue posible debido a la existencia de una matriz cuasi-singular. Posteriormente, el autor exploró la estimación de SUREs seleccionando algunas submuestras de ecuaciones y no se observaron cambios significativos en los resultados, en comparación a los ofrecidos por las estimaciones individuales. Estos cálculos están disponibles ante cualquier petición.

CUADRO 1  
**ESTIMACIÓN DE TASAS DE RENDIMIENTO AGREGADA EN EDUCACIÓN  
 POR REGIONES, EMPLEANDO SERIE TEMPORAL**

	cte	EM	ln (k)	R <sup>2</sup>	D.W.
Andalucía #	-0,5434 (3,03)*	0,0139 (0,70)	0,7526 (6,93)*	0,99	2,03
Aragón	-0,6869 (15,37)*	0,0514 (3,28)*	0,7483 (15,37)*	0,99	2,14
Asturias	-0,7132 (12,03)	-0,0249 (1,16)	0,9307 (13,11)*	0,99	1,56
Baleares	-0,4606 (4,96)*	0,1223 (3,68)*	0,5307 (6,14)*	0,96	1,54
Canarias	-0,7504 (11,66)*	0,0735 (2,76)*	0,7442 (9,57)*	0,98	1,90
Cantabria	-0,5569 (11,81)*	0,0522 (5,03)*	0,6326 (15,03)*	0,99	1,93
Cataluña	-0,2557 (6,70)*	0,0823 (7,52)*	0,5219 (12,92)*	0,99	2,18
Castilla-León	-0,5398 (16,61)*	0,0475 (3,74)*	0,6448 (21,44)*	0,99	2,18
Castilla-La Mancha #	-0,3719 (1,93)	0,0485 (1,76)	0,5608 (4,76)*	0,99	2,25
Comunidad Valenciana	-0,5755 (12,84)*	0,0543 (2,84)*	0,6878 (13,29)*	0,99	2,07
Extremadura	-0,7278 (14,12)*	0,0586 (2,34)*	0,6564 (13,70)*	0,99	1,82
Galicia	-0,8024 (24,28)*	0,0471 (2,67)*	0,7833 (20,64)*	0,99	2,14
Madrid	-0,4602 (12,31)*	0,0637 (6,15)*	0,7279 (15,42)*	0,99	1,93
Murcia	-0,8287 (11,53)*	-0,0109 (0,41)	0,9504 (11,48)*	0,98	1,46
Navarra	-0,4168 (6,49)*	0,0843 (3,70)*	0,5506 (7,00)*	0,98	1,60
País Vasco	-0,3035 (2,64)*	0,0449 (2,20)*	0,6088 (6,24)*	0,97	1,20
La Rioja	-0,2686 (6,05)*	0,0663 (4,02)*	0,5309 (15,39)*	0,99	2,30

Nota: El signo (\*) indica nivel de significación, al menos, el 5%. El signo (#) denota que la regresión ha sido ajustada por autocorrelación.

elevadas. En el otro extremo, regiones como Andalucía, Asturias, las dos Castillas o Galicia, usualmente identificadas como de escaso desarrollo, revelan tasas reducidas.

Una explicación posible de este patrón, sugerida por el propio Ram (1996), consistiría en el efecto supuestamente positivo que desempeña el dinamismo económico sobre las tasas de rendimiento. De tal manera que la rentabilidad será mayor en aquellas regiones donde el crecimiento económico sea más intenso, por cuanto éste hará más valorable las inversiones educativas<sup>23</sup>. Este efecto asociado al crecimiento podría llegar a compensar el advenimiento de los rendimientos decrecientes. Empíricamente, esta hipótesis puede ser investigada, a un nivel exploratorio, a través del siguiente modelo:

$$\ln(Y_{it}) = \alpha + \beta_1 EM_{it} + \beta_2 EM_{it} * DE_i + \gamma \ln(k_{it}) + \sum_{i=2}^n \delta_i D_i + \mu_{it} \quad (5)$$

El parámetro  $\beta_2$  recogería el efecto que sobre la tasa de rendimiento produce el dinamismo económico de la región. La variable  $DE_i$  se puede aproximar a través del crecimiento del PIB real de la región  $i$  en todo el período. Nótese que un signo positivo de  $\beta_2$  indicaría un apoyo a la hipótesis sobre los efectos beneficiosos de la vitalidad económica sobre las tasas de rendimiento educativo

El ajuste de la ecuación (5) produce los siguientes resultados:

$$\ln(Y) = -0,5825 + 0,0147 EM + 0,0234 EM*DE + 0,6577 \ln(k) + \dots,$$

(20,76)\* (1,32)                      (4,80)\*                      (33,42)\*

$$\bar{R}^2 = 0,98, \text{SEE} = 0,4167, \text{D.W.} = 1,95$$

En efecto, se confirma el efecto positivo que desempeña el dinamismo económico sobre las tasas de rendimiento de la educación. Así, las regiones con un mejor ambiente económico tienden a registrar tasas de rendimiento superiores, pudiendo contrarrestar los efectos negativos asociados a la ley de los rendimientos decrecientes.

23. Comentar, en concreto, que el grupo de regiones relativamente más desarrolladas ha registrado un crecimiento promedio en el período del 7.0%, mientras que las menos favorecidos han exhibido un registro del 6.4%

Finalmente, otro aspecto que puede ser analizado es la evolución en el tiempo de las tasas de rendimiento agregadas. Psacharopoulos (1989, 1994) ha ofrecido evidencia empírica que tiende a apoyar la hipótesis del declive general temporal de las tasas de rendimiento, en un contexto micro. Tiene interés analizar si este patrón temporal puede ser observado con los datos agregados regionales. Esta cuestión puede ser abordada empíricamente a través del empleo de las muestras cross-regionales. Como es lógico, en este análisis las variables ficticias no pueden ser utilizadas. Es oportuno señalar que el número de observaciones disponibles no es elevado, 17, por lo que el análisis ha de ser considerado con la evidente precaución. En cualquier caso, a un nivel simple e ilustrativo, parece útil realizar este ejercicio.

Los resultados correspondientes al ajuste de la ecuación (2), sin dummies, para cada una de las muestras cross-regionales se exhiben en el Cuadro 2. Algunos puntos merecen ser subrayados:

Primero, las tasas de rendimiento son positivas, elevadas y significativas en todos los años. Nótese que su magnitud es netamente superior a la desprendida con los datos de panel o los temporales. Esto sugiere que la educación tiene un efecto superior sobre la renta agregada cuando se analizan las diferencias entre regiones en un momento dado. No obstante, una posible sobreestimación puede ser verosímil a tenor de la omisión de las variables dummies.

Segundo, se constatan dos periodos claramente diferenciados en relación al patrón temporal de las tasas de rendimiento. Entre 1964 y 1973, periodo de elevado crecimiento económico y de aumentos educativos, la rentabilidad agregada registra una clara reducción, pasando del 48% al 18%. Esta observación se encuentra en línea con los hallazgos de Psacharopoulos. A partir de 1973, sin embargo, las tasas presentan oscilaciones, aunque manteniéndose entorno al 20%. Nótese que el patrón general es usualmente decreciente a excepción de los años en los que el ciclo económico se modifica: 1975 (inicio efectos 1ª crisis petrolera), 1981 (nueva crisis), 1987 (inicios expansión). Parecería, pues, que los cambios de ciclo producen un shock alcista en las tasas de rendimiento, al que le siguen posteriores decrecimientos hasta la aparición de un nuevo shock.

Finalmente, es instructivo hacer notar el alto poder explicativo del modelo cross-regional; en todos los años la capacidad de ajuste supera el 60%, llegando en algunos casos a aproximarse al 90%<sup>24</sup>.

24. No obstante, la bondad del ajuste es inferior a la obtenidas en los modelos de panel. Este resultado es lógico si se piensa que en los últimos se incluían variables dummies regionales.

CUADRO 2  
**ESTIMACIÓN DE TASAS DE RENDIMIENTO AGREGADA EN EDUCACIÓN  
 POR AÑOS, EMPLEANDO MUESTRAS CROSS-REGIONALES**

	cte	EM	ln (k)	R <sup>2</sup>	D.W.
1964	-1,7255 (6,93)*	0,4478 (4,48)*	0,4663 (2,76)*	0,80	1,78
1967	-1,6898 (11,34)*	0,4238 (6,57)*	0,5066 (3,68)*	0,87	2,02
1969	-1,5320 (8,66)	0,3347 (4,98)*	0,6333 (4,18)*	0,89	2,12
1971	-1,3806 (9,99)*	0,2531 (6,31)*	0,6988 (5,05)*	0,90	2,21
1973	-1,2024 (6,40)*	0,1758 (4,37)*	0,7826 (4,30)*	0,87	2,37
1975	-1,2074 (6,39)*	0,2209 (9,98)*	0,6400 (6,88)*	0,88	1,97
1977	-1,2075 (8,24)*	0,2025 (8,48)*	0,6600 (8,45)*	0,92	2,16
1979	-0,9179 (4,66)*	0,1798 (4,03)*	0,5590 (3,87)*	0,84	2,06
1981	-0,8703 (3,72)*	0,2007 (6,65)*	0,4782 (4,00)*	0,78	2,14
1983	-0,8892 (6,09)*	0,1914 (8,35)*	0,4872 (5,11)*	0,79	1,95
1985	-1,3429 (4,30)*	0,1875 (8,80)*	0,6750 (4,22)*	0,77	1,72
1987	-1,2940 (2,36)*	0,2031 (8,91)*	0,5942 (2,41)*	0,70	1,62
1989	-0,8250 (1,66)	0,1946 (7,87)*	0,3904 (1,76)	0,73	2,33
1991	-0,9289 (1,28)	0,1701 (4,43)*	0,4697 (1,38)	0,63	2,31

*Nota: El signo (\*) indica nivel de significación, al menos, el 5%. El signo (#) denota que la regresión ha sido ajustada por autocorrelación. Las estimaciones han sido corregidas por heteroscedasticidad utilizando el método de White (1980).*

---

#### 4. ALGUNOS COMENTARIOS SOBRE LA ESPECIFICACIÓN DEL MODELO

---

Tras el análisis de los resultados, algunas observaciones parecen especialmente apropiadas en relación a las especificaciones utilizadas. La principal justificación para la elección de las especificaciones es su sencillez e intuición y, sobretodo, su paralelismo con las ecuaciones de mincer estándares. Sin embargo, una mención explícita de dos aspectos parece conveniente.

Primero, dada la sencillez de los modelos empleados tiene interés evaluar si subsisten, por ejemplo, problemas significativos de variables omitidas. En este sentido, el estadístico de Jarque-Bera contrasta la hipótesis nula de residuos normalmente distribuidos, por lo que puede ser utilizado para juzgar de forma general la adecuación de los modelos<sup>25</sup>. Bajo la hipótesis nula el estadístico se comporta como una chi-cuadrado con dos grados de libertad. Pues bien, para el caso de la estimación básica (expresión (2)) el estadístico arroja un valor del 6,51, que pasa a ser del 2,76 para el caso no lineal (expresión (4)). El primer número supera el valor tabulado (5,99), mientras que el segundo no. En tales circunstancias, habríamos de concluir no sólo que el modelo básico presenta algún tipo de error sino que éste es únicamente atribuible a la necesidad de especificar una relación no lineal entre educación media y el logaritmo de la renta. Es preciso enfatizar, sin embargo, que, a pesar de esta observación, siguen teniendo un indudable interés analítico las estimaciones derivadas del modelo lineal por dos razones principales: 1) porque nos permiten calcular las tasas promedio de rendimiento y 2) porque nos permiten establecer un paralelismo con los conocidos modelos de Mincer.

Segundo, la cuestión concerniente a la endogeneidad potencial de la variable escolarización podría merecer algún comentario. Podría pensarse que la variable renta y la educación podrían estar mutuamente relacionadas, lo que podría implicar que el modelo uniecuacional utilizado fuera poco apropiado<sup>26</sup>. Un tratamiento adecuado de este aspecto requeriría la especificación de un modelo de ecuaciones simultáneas, que sobrepasa los límites y objetivos de este trabajo. No obstante, puede ser instructivo evaluar, a un nivel simple, si los errores de simultaneidad son relevantes en nuestras estimaciones unie-

25. Ramsey (1969) demostró que la omisión de variables relevantes, la especificación de una forma funcional incorrecta y la existencia de correlación entre las variables explicativas y los residuos producen vectores de residuos con media no nula y, por tanto, no normales.

26. Es oportuno señalar que no está demasiado clara la racionalidad teórica acerca del por qué el output real corriente debe afectar al stock educativo corriente.

cuacionales. En este sentido, el test de Jarque-Bera vuelve a ser muy útil, por cuanto una de las principales consecuencias de los errores de simultaneidad está asociada, precisamente, con la generación de residuos no normales. Dado que los residuos no normales de la ecuación básica se achacan a la relación no lineal que vincula educación con renta y dado que la ecuación cuadrática produce perturbaciones normales habríamos de inferir que los sesgos por simultaneidad no son un problema relevante en nuestras estimaciones.

---

## 5. CONSIDERACIONES FINALES

---

Estudiar las tasas de rendimiento de la educación tiene una obvia relevancia. Este artículo presenta una primera evidencia del estudio de este aspecto en un marco agregado regional. Aunque los modelos utilizados seguro no son perfectos y los datos habrían de ser mejorados, parece útil un análisis como el realizado, por cuanto éste parece razonable, intuitivo y arroja algunas ventajas operativas apreciables.

Los consideraciones más interesantes que pueden desprenderse del trabajo podrían ser las siguientes:

Primero, la tasa agregada promedio de rendimiento de la educación, estimada a través de datos multitemporales regionales, se sitúa cercana al 6%. Este valor está en línea con los resultados encontrados por otros autores. Esta cifra tiene la ventaja de incorporar los posibles efectos externos asociados a las inversiones educativas, generalmente no considerados en el resto de estudios.

Segundo, a través de una especificación exploratoria se ha encontrado evidencia significativa en favor de la existencia de una relación no lineal entre las inversiones educativas y su rentabilidad. En particular, parece que el producto marginal agregado de las expansiones educativas es decreciente. De esta forma, cabría esperar que los efectos beneficiosos de las inversiones educativas, *ceteris paribus*, se vayan reduciendo a medida que avanza el nivel de instrucción.

Tercero, los resultados revelan que las regiones de desarrollo relativo medio y elevado presentan en promedio una tasa de rendimiento superior a la exhibida por las regiones más atrasadas. Una posible hipótesis explicativa de este resultado se basa el efecto beneficioso del crecimiento económico, el cual ha sido mayor en las primeras. Un modelo tentativo ha confirmado empíricamente esta intuición.

Cuarto, las tasas de rendimiento agregadas parecen haber declinado notablemente hasta 1973, fecha a partir de la cual sufren continuas oscilaciones. Podría especularse que esta última variabilidad podría estar vinculada a efectos cíclicos.

Quinto, se ha empleado un test estadístico para contrastar empíricamente la validez de las especificaciones utilizadas. El test Jarque-Bera indica la ausencia de errores relevantes en el modelo no lineal. Ello confirma que los problemas empíricos del modelo básico se atribuyen única y exclusivamente a la idoneidad de postular rendimientos decrecientes en la educación. No obstante, sigue teniendo un obvio interés los resultados desprendidos del modelo básico, cuyo paralelismo con los modelos de mincer es atractivo, y que permite recabar información sobre las tasas promedio de rendimiento.

Sexto, sería interesante verificar los patrones encontrados en el trabajo con otros datos. En particular, parece útil reevaluar el análisis con la ayuda de datos educativos cualitativamente mejores a los aquí empleados.

## BIBLIOGRAFÍA

- ALBA, A. y M.J. SAN SEGUNDO, (1995), "The Returns to Education in Spain", *Economics of Education Review*, Vol. 14, nº 2, págs. 155-166.
- BECKER, G., (1964), *Human Capital*, National Bureau of Economic Research.
- DURO, J.A., (1997), "Expansión educativa y desigualdad en los niveles de escolarización: un test empírico para España", *Información Comercial Española. Revista de Economía*, nº764, Julio-Agosto, págs. 147-160
- FUNDACION BBV, (1996), *El stock de capital en España y sus Comunidades Autónomas*, tres volúmenes.
- GARCIA CALVO, J.L., (1988), "Rendimientos del capital humano en educación en España", *Investigaciones Económicas*, Vol. XII, nº 3, págs 473-482.
- GRILICHES, Z, (1977), "Estimating the Returns to Schooling: Some Econometric Problems", *Econometrica*, 45, págs. 1-22.
- LASSIBILLE, G., (1988), "La formación de las rentas del trabajo en Andalucía", *Investigaciones Económicas*, Vol. 12, nº3, págs. 483-499.
- MANKIW, G, D. ROMER y D. WEIL, (1992), "A contribution to the empirics of economic growth", *Quarterly Journal of Economics*, Vol. 107, nº 2, págs. 407-437.
- MAS, M, F. PEREZ, E. URIEL y L. SERRANO, (1995), *Capital Humano. Series Históricas, 1964-1992*, Fundación Bancaja
- MINCER, J., (1974), *Schooling, Experience, and Earnings*, NBER, N.Y.
- PSACHAROPOULOS, G., (1980), "Returns to education: an updated international comparison" en *Education and Income*, T. King (ed.), Staff Working Paper nº 402, Washington-The World Bank
- PSACHAROPOULOS, G., (1989), "Time Trends of the Returns to Education: Cross-National Evidence", *Economics of Education Review*, Vol. 8, nº 3, págs. 225-231.
- PSACHAROPOULOS, G., (1994)", "Returns to Investment in Education: A Global Update", *World Development*, Vol.22, nº 9, págs. 1325-1343.
- RAM, R., (1996), "Level of Development and Rates of Return to Schooling: Some Estimates from Multicountry Data", *Economic Development and Cultural Change*, Vol. 44, nº 4, págs. 839-857.
- RAMSEY, J., (1969), "Tests for Specification Errors in Classical Linear Least Squares Regression Analysis", *Journal of the Royal Statistical Society*, Series B, Vol. 31, págs. 350-371.
- SOLOW, R., (1958), "A contribution to the Theory of Economic Growth", *Quarterly Journal of Economics*, 70, págs. 65-94.

- SCHULTZ, T.P., (1988), "Education Investments and Returns", en *Handbook of Development Economics*, H. Chenery y T.N. Srinivasan (eds), Vol, I, North Holland.
- SCHULTZ, T.W., (1961), "Investment in Human Capital", *American Economic Review*, págs. 1-17.
- WHITE, H., (1980), "A Heteroskedasticity-Consistent Covariance Estimator and a Direct Test for Heteroskedasticity", *Econometrica*, Vol. 48, págs. 817-838.

Recibido, 9 de Octubre de 1997; Aceptado, 9 de Enero de 1998.