

# **Desajuste educativo y salarios en España: nueva evidencia con datos de panel(\*)**

por

M<sup>a</sup> ISABEL AGUILAR RAMOS  
DOLORES GARCÍA CRESPO

Departamento de Economía Aplicada. Universidad de Málaga

## **RESUMEN**

Este trabajo cuantifica el desajuste educativo en el mercado laboral español y sus efectos sobre los salarios de los trabajadores a partir del PHOGUE (1995-2001). Las estimaciones por MCO de la función de salarios indican que los individuos con desajuste educativo sufren una penalización salarial en relación con aquellos que poseen un nivel educativo acorde con el puesto que ocupan. Sin embargo, al controlar por la heterogeneidad individual inobservada mediante un modelo de efectos aleatorios, los coeficientes estimados de las variables de desajuste educativo se reducen de forma significativa, cuestionando los resultados tradicionales obtenidos con muestras de corte transversal.

---

(\*) Las autoras agradecen las sugerencias recibidas de los participantes en las XII Jornadas de Economía de la Educación y en el VII Encuentro de Economía Aplicada, así como los comentarios de dos evaluadores anónimos que han permitido mejorar sustancialmente el contenido final del trabajo. Además, agradecen la financiación obtenida del proyecto SEC2003-08855-C01-01 del Ministerio de Educación y Ciencia.

*Palabras clave:* desajuste educativo, salarios, panel de datos.

*Clasificación AMS:* 62P20, 91B40.

## 1. INTRODUCCIÓN

Las investigaciones sobre el mercado laboral muestran que una determinada proporción de la población ocupada presenta un desajuste entre los estudios poseídos y los requeridos por los empleos, bien por exceso (sobreeducación) o por defecto (infraeducación). Inicialmente este tema despertó el interés de los investigadores estadounidenses en los años setenta, y posteriormente se extendió a los países europeos, y de forma especial, al Reino Unido, Holanda, España o Portugal (Groot y Maassen van den Brink, 2000; Hartog, 2000 y Aguilar, 2005 constituyen revisiones de estas aportaciones).

Junto a la cuantificación del desajuste educativo en la población ocupada, un segundo objetivo de algunos de estos estudios ha sido evaluar el impacto de tal desajuste sobre los salarios percibidos por los trabajadores. En términos generales, y con datos de corte transversal, las conclusiones obtenidas han sido que la rentabilidad de los años de sobreeducación es positiva, aunque menor a la de los años de educación requerida por los empleos, mientras que la rentabilidad de los años de infraeducación suele ser negativa (Hartog, 2000). Sin embargo, en España existen aún pocas investigaciones sobre este tema debido a la falta de información adecuada (véanse Alba, 1993; Beneito *et al.*, 1996 y Alba y Blázquez, 2004).

Recientemente, y en una línea distinta a la anterior, han aparecido algunos trabajos que llaman la atención sobre una posible sobreestimación de la incidencia del desajuste educativo en el mercado laboral, así como del efecto negativo que la sobreeducación ejerce sobre los salarios de los individuos. Así, los trabajos de Chevalier (2003), Robst (1995) y Walker y Zhu (2005) para el Reino Unido y Gautier *et al.* (2002) para Holanda descansan en la idea de que el colectivo de trabajadores sobreeducados resulta muy heterogéneo, por lo que cualquier análisis que intente cuantificar la posible penalización salarial asociada a un exceso de años de estudios debe desarrollar una metodología apropiada que tenga en cuenta tal heterogeneidad, tras lo cual, y de acuerdo con la evidencia empírica existente, los resultados cambian de forma significativa.

Así, la penalización salarial que habitualmente se ha atribuido a los años de sobreeducación de un trabajador, en comparación con el individuo que realiza un trabajo con el nivel de estudios apropiado al que posee, en realidad podría estar

recogiendo otros efectos tales como la calidad de la universidad (Robst, 1995), las características del emparejamiento trabajador-empresa (Gautier *et al.*, 2002) o la capacidad innata del individuo (Chevalier, 2003 y Walker y Zhu, 2005), entre otras. En definitiva, este tipo de características individuales deberían estar presentes en las ecuaciones salariales especificadas, pero el problema para llevar a cabo este planteamiento es la dificultad de observarlas o incluso de medirlas adecuadamente.

Ante este problema habitual de ausencia de información necesaria para completar las ecuaciones de salarios, Bauer (2002) propone una solución econométrica que consiste en utilizar bases de datos de tipo longitudinal para analizar el comportamiento de los agentes teniendo en cuenta la heterogeneidad individual inobservada. Bajo este enfoque alternativo, la estrategia consiste en utilizar técnicas de estimación adecuadas para superar las limitaciones de los estudios empíricos que abordan el tema con muestras de corte transversal. Para el mercado laboral alemán, las nuevas estimaciones muestran una reducción importante en la penalización salarial asociada al desajuste educativo obtenida hasta entonces con datos trasversales.

Siguiendo esta línea, la disponibilidad actual de una fuente estadística longitudinal como es el Panel de Hogares de la Unión Europea (PHOGUE) nos va a permitir estimar las habituales funciones salariales utilizando la estructura de panel de datos, comprobando de esta forma si para el mercado laboral español se mantienen los resultados tradicionales respecto a la rentabilidad monetaria de la educación o, al igual que en Alemania, se reduce significativamente la penalización asociada al estado de sobreeducado.

El artículo presenta la siguiente estructura. En el segundo apartado se realiza una breve descripción de los datos y variables utilizados, prestando una atención especial a la medida propuesta para valorar el desajuste educativo y a la distribución de los ingresos. En el siguiente apartado se discute la metodología econométrica utilizada para analizar el efecto del desajuste educativo sobre los ingresos de los trabajadores. A continuación, en el cuarto epígrafe se presentan los resultados obtenidos y, por último, en el quinto se resumen las principales conclusiones extraídas.

## **2. DATOS Y VARIABLES**

La fuente de datos utilizada es el PHOGUE, una encuesta longitudinal de periodicidad anual que se realiza de forma estandarizada en la Unión Europea y ofrece información sobre un amplio espectro de características de la población (demogra-

fía, oferta de trabajo, renta, salud, educación, etc.)(1). En nuestro país, el Instituto Nacional de Estadística recoge esta información a partir de una muestra representativa de la población española, compuesta aproximadamente por 17.000 individuos y 6.500 hogares.

En este trabajo se va a considerar la información correspondiente al periodo 1995-2001(2). Siguiendo el criterio utilizado habitualmente en la literatura, en cada uno de estos años se han seleccionado a los trabajadores asalariados con edades comprendidas entre los 16 y los 60 años, excluyendo a los ocupados en agricultura, así como a los que aún continúan estudiando. Tras eliminar aquellas observaciones cuya información para algunas variables relevantes no se encuentra disponible, la submuestra final con la que se ha llevado a cabo el análisis descriptivo y econométrico se eleva a 6.112 individuos que forman un panel incompleto de 18.223 observaciones durante las siete olas consideradas(3).

Dado que las variables más interesantes de cara a cubrir los objetivos planteados son aquellas que permiten comparar el nivel educativo de los trabajadores con la ocupación que desempeñan, así como las que recogen los ingresos salariales percibidos, a continuación se presenta una breve descripción de las mismas.

No existe un método unánimemente aceptado para medir el desajuste educativo de los trabajadores, ya que cada una de las medidas propuestas en la literatura tiene ventajas e inconvenientes. En cualquier caso, las principales medidas utilizadas pueden clasificarse en tres grandes grupos: objetivas, subjetivas y estadísticas.

En primer lugar, las medidas objetivas parten de la evaluación de los puestos de trabajo por el propio investigador, por algún organismo oficial o por expertos, y de la clasificación de los mismos en función del nivel educativo que se considera necesario para su realización. Este nivel se compara con la formación de los individuos para ver si es o no la adecuada(4). La objetividad de este análisis sistemático de los empleos podría convertir a este método en uno de los más atractivos para medir el

---

(1) Véase Peracchi (2002) para una descripción completa del Panel de Hogares de la Unión Europea.

(2) No se incluye la primera ola de la encuesta porque en ella no aparecen algunas variables relevantes, como el tipo de contrato, y porque en algunos casos la codificación difiere de la de los ciclos posteriores.

(3) Desde el punto de vista econométrico, el hecho de que el panel sea incompleto sólo supone un mayor coste computacional en el cálculo de la matriz de varianzas y covarianzas (Greene, 1998).

(4) El *U.S. Dictionary of Occupational Titles*, que proporciona información sobre la educación requerida por un amplio conjunto de ocupaciones, constituye un claro referente, ya que se ha utilizado en la mayoría de los trabajos realizados en Estados Unidos para medir el desajuste educativo.

desajuste educativo. Sin embargo, resulta difícil de aplicar porque la tarea de clasificación de los puestos es bastante compleja y precisa gran cantidad de información. Además, la evaluación externa está sujeta a cierto margen de error, al basarse en clasificaciones necesariamente amplias que no siempre pueden recoger rasgos específicos de los empleos ni los cambios en los niveles educativos requeridos derivados del avance tecnológico.

Por su parte, las medidas subjetivas se basan en la información suministrada por los propios trabajadores. Una primera posibilidad sería preguntarles directamente si consideran que la educación que poseen es la adecuada para desempeñar su trabajo o si, por el contrario, creen que están sobreeducados o infraeducados. Una segunda opción, que podríamos llamar indirecta, consistiría en comparar el nivel educativo del trabajador con el que éste declara óptimo o necesario para realizar su trabajo correctamente. Si se tiene en cuenta que una determinada ocupación puede presentar importantes diferencias según la región, la empresa o el tipo de industria en la que se lleva a cabo, podría pensarse que las definiciones subjetivas resultan más adecuadas que las objetivas, ya que el trabajador conocerá mejor su puesto que cualquier investigador u organismo oficial y podrá evaluarlo con más precisión. Además, estos métodos resultan menos costosos porque requieren menos información. Sin embargo, la validez de los indicadores subjetivos también se ha cuestionado a menudo, no sólo por la carga sentimental que encierran, sino porque exigen a los trabajadores algo que no es fácil, evaluar la formación necesaria para desempeñar una función.

Por último, las medidas estadísticas calculan la educación requerida con base al nivel educativo promedio de las ocupaciones. Verdugo y Verdugo (1989) proponen considerar sobrecualificados a los trabajadores cuyos años de educación formal superan en una desviación típica a los años medios de su ocupación, e infracualificados a aquellos que se sitúan por debajo de una desviación típica de la media. Aunque este método resulta atractivo porque necesita muy poca información, también ha sido criticado. En primer lugar, porque se acepta implícitamente que los años adecuados de educación se sitúan en torno a la media, cuando pueden existir ocupaciones en las que predomine la sobreeducación o la infraeducación. En segundo lugar, porque la elección de una desviación estándar como punto de corte parece arbitraria. Y en tercer lugar, porque la clasificación en ocupaciones de dos o tres dígitos puede ser demasiado burda. En la práctica, esta medida suele arrojar menores desajustes que las definiciones objetivas o subjetivas. Además, suelen encontrarse porcentajes similares de trabajadores infra y sobrecualificados (en muchos casos, cercanos al 16%) por lo que algunos autores han argumentado que

estos resultados indican simplemente que los años de educación de los trabajadores en una ocupación determinada siguen una distribución normal(5).

En este trabajo se va a proponer una medida subjetiva del desajuste educativo, elaborada a partir de dos preguntas incluidas en el apartado de empleo del cuestionario del PHOGUE(6). En primer lugar, se formula a los trabajadores la siguiente pregunta (PE016): “¿considera que sus conocimientos o sus condiciones personales le permitirían realizar un trabajo más cualificado?”. Esta información se complementa con la siguiente cuestión (PE021): “sus estudios o su formación, ¿le han proporcionado los conocimientos necesarios para desempeñar su trabajo actual?”.

Como muestra la tabla 1, la primera pregunta permite diferenciar dos categorías de trabajadores: los sobreeducados (aquellos que contestan que sí podrían realizar un trabajo más cualificado) y los no sobreeducados (los que contestan que no). A continuación, en base a la pregunta PE021 pueden diferenciarse distintos grupos dentro de estas categorías. En primer lugar, se distingue entre aquellos trabajadores sobreeducados cuyos estudios les han proporcionado los conocimientos necesarios para desempeñar su trabajo (sobreeducados tipo 1) y los que no han recibido estos conocimientos gracias a los estudios o la formación (sobreeducados tipo 2). De forma paralela, dentro del colectivo no sobreeducado, si el trabajador indica que sus estudios le han proporcionado los conocimientos necesarios para desarrollar su trabajo se considera adecuadamente educado y en caso contrario, infraeducado.

---

(5) Alternativamente, Kiker *et al.* (1997) sugieren tomar como referencia la moda en lugar de la media. De esta forma, los trabajadores cuyo nivel educativo coincide con el valor modal para su ocupación se consideran adecuadamente educados, los que se superan ese valor, sobreeducados y los que se encuentran por debajo, infraeducados.

(6) La información del PHOGUE resulta limitada para plantear medidas estadísticas u objetivas, ya que las ocupaciones de los trabajadores se codifican de acuerdo con la clasificación internacional ISCO88 a un nivel de dos dígitos, agrupándose además algunos epígrafes de la misma.

**Tabla 1**  
**CLASIFICACIÓN DE LOS OCUPADOS SEGÚN DESAJUSTE EDUCATIVO A PARTIR DEL PHOGHE**

		Pregunta <i>PE021</i>			
		Sí	No		
		↓	↓		
Pregunta <i>PE016</i>	→ Sí	Sobreeducados tipo 1	Sobreeducados tipo 2	→ Sobreeducados	
	→ No	Adecuadamente educados	Infraeducados	→ No sobreeducado	

Nota: la pregunta *PE016* establece "¿considera que sus conocimientos o sus condiciones personales le permitirían realizar un trabajo más cualificado?" y la *PE021* "sus estudios o su formación, ¿le han proporcionado los conocimientos necesarios para desempeñar su trabajo actual?".

La tabla 2 recoge la distribución de los individuos ocupados de acuerdo con esta medida de desajuste educativo en tres de los años considerados. Estas cifras reflejan, en primer lugar, una elevada incidencia del desajuste educativo (sobre todo de la sobreeducación) en el mercado de trabajo. Así, la mayoría de los asalariados declara que podría realizar un trabajo más cualificado, de manera que el porcentaje de sobreeducados supera el 50% en todos los casos. En segundo lugar, se aprecia que, ante el continuo incremento de los niveles educativos de la población activa en las últimas décadas, se produce un descenso de la proporción de trabajadores infraeducados a lo largo del periodo considerado, junto con un incremento del porcentaje de sobreeducados. En tercer lugar, se detectan importantes diferencias por género, siendo la sobreeducación (al menos, la de tipo 1) más frecuente entre las mujeres y la infraeducación, entre los varones, en línea con los resultados obtenidos en García Serrano y Malo (1997).

**Tabla 2.**  
DESAJUSTE EDUCATIVO SEGÚN SEXO (%)

	1995		
	<i>Total</i>	<i>Varones</i>	<i>Mujeres</i>
Sobreeducados Tipo 1	31,36	28,37	36,93
Sobreeducados Tipo 2	22,06	23,25	19,85
Adecuadamente educados	21,54	21,01	22,53
Infraeducados	25,03	27,37	20,68
Número de observaciones	3.096	2.013	1.083

	1998		
	<i>Total</i>	<i>Varones</i>	<i>Mujeres</i>
Sobreeducados Tipo 1	34,75	32,55	37,95
Sobreeducados Tipo 2	23,54	24,44	22,03
Adecuadamente educados	21,70	20,07	24,43
Infraeducados	20,19	22,94	15,59
Número de observaciones	2.447	1.530	917

	2001		
	<i>Total</i>	<i>Varones</i>	<i>Mujeres</i>
Sobreeducados Tipo 1	35,00	32,24	39,26
Sobreeducados Tipo 2	24,93	24,24	24,44
Adecuadamente educados	19,14	17,35	21,90
Infraeducados	20,93	25,17	14,39
Número de observaciones	2.403	1.458	945

Fuente: elaboración propia a partir del PHOGUE (1995-2001).

Con objeto de comparar estos resultados con los obtenidos en otras investigaciones también referidas a la población española, en la tabla 3 se recogen algunos de los trabajos más destacados. Lógicamente, la comparación ha de realizarse con la debida cautela, no sólo porque estos estudios utilizan distintos procedimientos para medir el desajuste educativo, sino porque además se han realizado en diversos momentos temporales, con información procedente de distintas encuestas.



**Tabla 3**  
MEDIDA DEL DESAJUSTE EDUCATIVO EN ESPAÑA SEGÚN DIVERSOS ESTUDIOS (%)

<i>Autor</i>	<i>Encuesta (año)</i>	<i>Medida del desajuste</i>	<i>Sobreeducados</i>
Alba (1993)	ECVT (1985)	Subjetiva (indirecta)	17,1
Alba y Blázquez (2004)	PHOGUE (1995)	Subjetiva (directa)	53,8
Beneito <i>et al.</i> (1996)	ECBC (1991)	Subjetiva (directa)	27,9
		Subjetiva (indirecta)	25,6
		Estadística	15,2
García Montalvo (1995)	EPA (1985)	Objetiva	3,7
		Objetiva	5,1
		Objetiva	6,3
		Objetiva	6,6
		Objetiva	7,7
		Estadística	8,9
García Serrano y Malo (1996, 1997)	ECBC (1991)	Subjetiva (directa)	29,9 (26,0/36,6)*
		Subjetiva (indirecta)	29,6 (27,0/34,1)*

<i>Autor</i>	<i>Adecuados</i>	<i>Infraeducados</i>
Alba (1993)	59,7	23,1
Alba y Blázquez (2004)	24,0	22,2
Beneito <i>et al.</i> (1996)	61,3	10,9
	58,0	16,5
	69,4	15,3
García Montalvo (1995)	65,9	30,4
	63,7	31,2
	62,6	31,1
	62,9	30,5
	64,7	27,6
	84,9	6,2
García Serrano y Malo (1996, 1997)	59,5	10,6
	(39,5/43,9)*	(34,5/19,5)*
	41,7 (61,1/57,7)*	28,7 (11,9/8,2)*

(\*) (varones/mujeres).

Fuente: elaboración propia a partir de los trabajos citados.

En cualquier caso, puede comprobarse que, si bien las medidas subjetivas arrojan siempre los mayores porcentajes de sobreeducación, las cifras obtenidas en nuestro trabajo superan a las estimadas tradicionalmente, con excepción de las que aparecen en Alba y Blázquez (2004), tal y como se podía esperar, dado que también se basan en la información del PHOGUE(7).

A la hora de valorar esta discrepancia, resulta difícil pensar que el resultado obtenido refleje simplemente un incremento temporal de la sobreeducación. En realidad, debe tenerse en cuenta que la medida de sobreeducación construida a partir de la información del Panel resulta más amplia que la considerada en otros trabajos. Así, de acuerdo con la pregunta PE016 presentada anteriormente, cuando un individuo evalúa si podría realizar un trabajo más cualificado no sólo debe tener en cuenta sus conocimientos sino también sus propias "condiciones personales". Sin embargo, las medidas subjetivas recogidas en Alba (1993), Beneito *et al.* (1996) o García Serrano y Malo (1996, 1997) se obtienen a partir de preguntas referidas sólo a los estudios de los trabajadores y no a otras características personales de los mismos(8).

No obstante, como se ha indicado anteriormente, esta clasificación inicial se ha depurado a partir de la cuestión PE021, distinguiéndose dos tipos de sobreeducados. Así, los que indican que sus estudios o su formación les han proporcionado los conocimientos necesarios para desempeñar su trabajo (tipo 1) pueden ser considerados sobreeducados en un sentido más estricto, en línea con el concepto utilizado normalmente en la literatura. Sin embargo, los sobreeducados tipo 2 constituyen un grupo heterogéneo de ocupados, ya que parece que su desajuste no se deriva de la formación recibida sino más bien de otras condiciones personales, tal y como indica el hecho de que el 68% de estos trabajadores tengan estudios primarios o inferiores y sólo el 10% estudios universitarios, según se muestra en la tabla 4. De esta forma, este grupo podría calificarse como "aparentemente sobreeducado" y

---

(7) Aunque estos autores utilizan el PHOGUE como fuente estadística, consideran una pregunta adicional del cuestionario para medir el desajuste educativo (PE022: "¿en qué medida está su trabajo actual relacionado con su formación?").

(8) En concreto, la ECVT incluye la siguiente cuestión: "¿qué clase de preparación necesita una persona para tener el trabajo que usted desempeña?", mientras que en la ECBC aparecen dos preguntas que permiten construir dos medidas alternativas del desajuste educativo: "¿diría usted que sus estudios son más que suficientes, suficientes o insuficientes para su trabajo actual?" y "¿qué nivel de estudios cree usted que es, actualmente, el más adecuado para realizar el trabajo que usted desempeña?". En esta última encuesta también se formula una pregunta más general que recoge el desajuste en la preparación y no sólo el referido al nivel de estudios: "con la experiencia que tiene ahora ¿diría usted que su preparación es más que suficiente, suficiente o insuficiente?". García Serrano y Malo (1996) indican que con esta pregunta obtienen resultados similares a los recogidos en la tabla 3, aunque no los presentan.

debería analizarse con suma cautela. Esta sospecha se ve reforzada al observar las funciones de densidad de los salarios de los cuatro grupos de trabajadores considerados de acuerdo con su desajuste educativo.

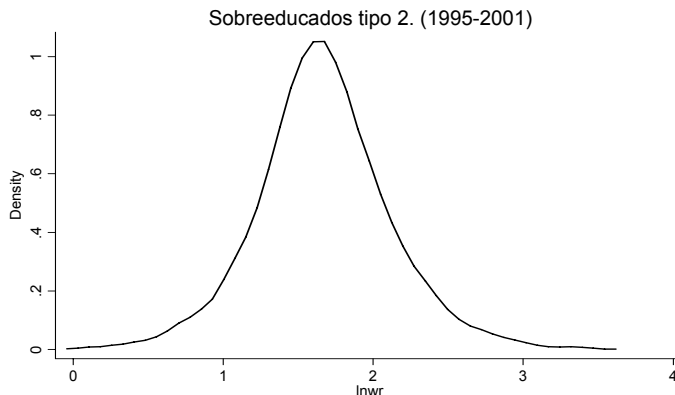
**Tabla 4**  
NIVEL DE ESTUDIOS SEGÚN DESAJUSTE EDUCATIVO DE LOS TRABAJADORES (%)

	<i>Sobreeducados Tipo 1</i>	<i>Sobreeducados Tipo 2</i>	<i>Adecuadamente educados</i>	<i>Infraeducados</i>
E. primarios o inferiores	33,18	68,42	45,53	88,90
E. medios	27,49	20,88	20,09	8,52
E. universitarios	39,32	10,69	34,36	2,50
Número de observaciones	6.088	4.272	3.896	3.967

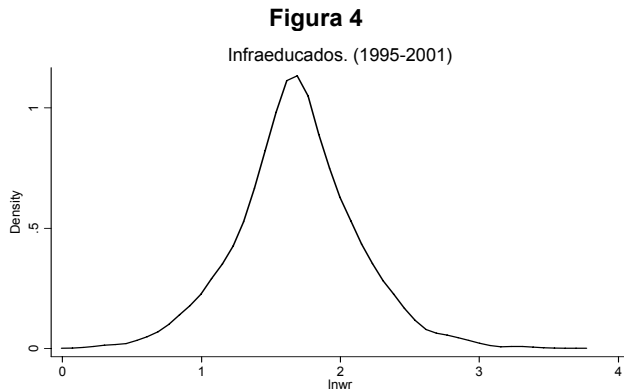
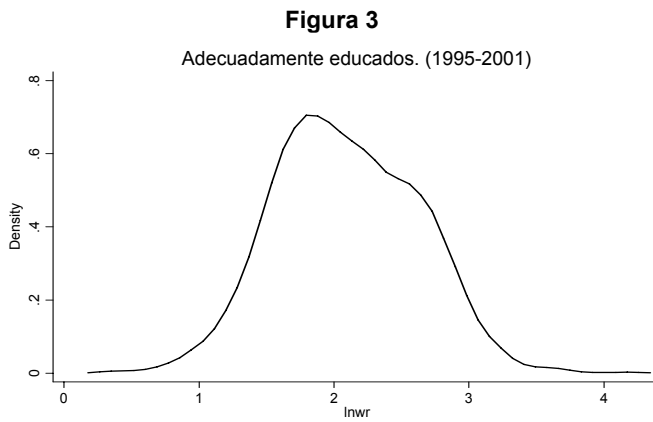
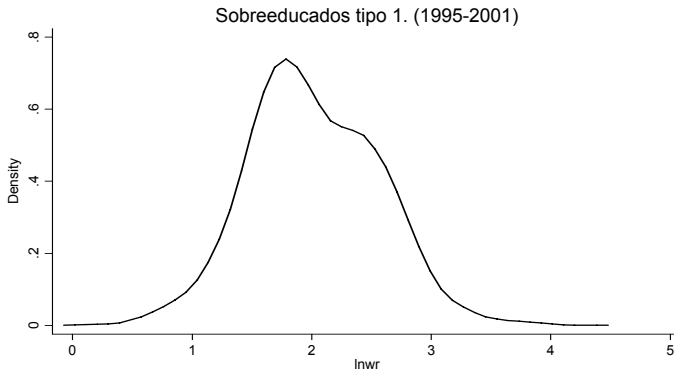
Fuente: elaboración propia a partir del PHOGUE (1995-2001).

Las Figuras 1-4 muestran las estimaciones Kernel de estas funciones y en ellas puede observarse la gran similitud existente entre las correspondientes a los individuos sobreeducados tipo 2 y los infraeducados.

**Figura 1**  
ESTIMACIONES KERNEL DE LA FUNCIÓN DE DENSIDAD DEL LOGARITMO DEL SALARIO HORA BRUTO SEGÚN DESAJUSTE EDUCATIVO



**Figura 2**  
ESTIMACIONES KERNEL DE LA FUNCIÓN DE DENSIDAD DEL LOGARITMO  
DEL SALARIO HORA BRUTO SEGÚN DESAJUSTE EDUCATIVO



Para finalizar este apartado descriptivo, la tabla 5 presenta el salario hora bruto según desajuste educativo y nivel de estudios de los individuos. En ella destacan tres aspectos: (i) en primer lugar, el salario medio de los universitarios duplica al de los que tienen estudios primarios o inferiores, independientemente del desajuste educativo que presenten (a excepción del grupo de sobreeducados tipo 2, donde la diferencia se reduce ligeramente); (ii) los salarios más elevados se corresponden con los individuos sin desajuste educativo; (iii) existe una penalización salarial asociada al desajuste, de mayor cuantía para los sobreeducados tipo 2 y los infraeducados.

**Tabla 5**

SALARIO HORA BRUTO SEGÚN DESAJUSTE EDUCATIVO Y NIVEL DE ESTUDIOS (EUROS DE 2001)

	<i>Sobreeducados Tipo 1</i>	<i>Sobreeducados Tipo 2</i>	<i>Adecuadamente educados</i>
E. primarios o inferiores	6,03 (-0,52)	5,37 (-1,18)	6,55
E. medios	7,44 (-1,59)	6,15 (-2,88)	9,03
E. universitarios	12,11 (-1,08)	8,46 (-4,73)	13,19
Total	8,81 (-0,52)	5,87 (-3,46)	9,33

	<i>Infraeducados</i>	<i>Total</i>
E. primarios o inferiores	5,65 (-0,90)	5,80
E. medios	7,38 (-1,65)	7,46
E. universitarios	12,52 (-0,67)	12,07
Total	5,97 (-3,36)	7,61

Nota: penalización salarial respecto a los adecuadamente educados entre paréntesis (euros de 2001).

Fuente: elaboración propia a partir del PHOGUE (1995-2001).

Aunque a partir de estas cifras no pueden obtenerse resultados concluyentes respecto al rendimiento salarial de la educación en estos colectivos, sin duda nos ofrecen una primera visión de las similitudes o diferencias entre los ingresos de los distintos tipos de trabajadores. Así, tenemos en mente que *a priori* el desajuste educativo provoca en general una penalización salarial en todas las categorías

educativas. Además, el colectivo de sobreeducados de tipo 2 es muy particular y presenta una densidad salarial muy similar a la de los infraeducados.

A continuación, en el siguiente epígrafe abordamos el análisis econométrico que nos permitirá comprobar si al tener en cuenta la heterogeneidad individual se mantiene la penalización salarial inicialmente observada. La descripción estadística de las variables incluidas en las estimaciones se presenta en el Apéndice.

### 3. ESPECIFICACIÓN ECONOMÉTRICA

Para analizar la relación entre educación e ingresos partimos de las siguientes ecuaciones salariales:

$$\ln Y_{it} = \alpha + \beta_1 E_{it} + \beta_2 \text{Exp}_{it} + \beta_3 \text{Exp}_{it}^2 + u_{it} \quad [1]$$

$$\ln Y_{it} = \alpha + \beta_1 E_{it} + \gamma_1 \text{Sobre1}_{it} + \gamma_2 \text{Sobre2}_{it} + \gamma_3 \text{Infra}_{it} + \beta_2 \text{Exp}_{it} + \beta_3 \text{Exp}_{it}^2 + u_{it} \quad [2]$$

$$\ln Y_{it} = \alpha + \beta_1 E_{it} + \gamma_1 \text{Sobre1}_{it} + \gamma_2 \text{Sobre2}_{it} + \gamma_3 \text{Infra}_{it} + \beta_2 \text{Exp}_{it} + \beta_3 \text{Exp}_{it}^2 + \beta_4 X_{it} + u_{it} \quad [3]$$

donde  $\ln Y_{it}$  hace referencia al logaritmo del salario hora bruto del individuo  $i=1, \dots, n$  en el periodo  $t=1995, \dots, 2001$ ;  $E_{it}$  representa los años de educación(9),  $\text{Exp}_{it}$  denota la experiencia en el mercado de trabajo; *Sobre1*, *Sobre2* e *Infra* son tres variables ficticias que toman el valor 1 si el individuo está sobreeducado tipo 1, tipo 2 o infraeducado, respectivamente;  $X_{it}$  es un vector que contiene otras características explicativas de los salarios individuales (tanto personales, como del puesto de trabajo, de la historia laboral o del entorno económico) y  $u_{it}$  es un término de error.

En la ecuación [1], la más parsimoniosa de todas, el coeficiente  $\beta_1$  mide el rendimiento de la educación en un sentido amplio, ya que toda la influencia indirecta de la educación sobre los salarios, como la elección de ocupación o el sector, se le

---

(9) Para obtener la variable de educación en forma continua, se ha asignado al máximo nivel educativo de los individuos un número teórico de años de estudios. En concreto, 2 años a los estudios inferiores a los Primarios, 5 a los Estudios Primarios, 8 al Primer nivel de la Enseñanza Secundaria, 9 a la Formación Profesional de Primer Grado, 11 a la Formación Profesional de Segundo Grado y FP3, 12 al Segundo Nivel de Enseñanza Secundaria, 15 a las Diplomaturas y 17 a las Licenciaturas y Doctorados.

atribuye directamente a ésta (Fersterer y Winter-Ebmer, 2003). La ecuación [2] permite analizar los efectos del desajuste educativo en los salarios (Verdugo y Verdugo, 1989). Para ello, los trabajadores con desajuste educativo se comparan con aquellos que poseen el mismo nivel de estudios y la misma experiencia, pero que ocupan puestos que se corresponden con su formación, es decir, se encuentran adecuadamente educados.

En esta especificación, si la productividad y los salarios se determinan en función del nivel educativo del trabajador, como postula la teoría del capital humano, entonces  $\gamma_1 = \gamma_2 = \gamma_3 = 0$ . Sin embargo, si de acuerdo con la teoría de la competencia por los puestos de trabajo (Thurrow, 1975), los salarios dependen del nivel educativo requerido por los empleos, los años de sobreeducación serían improductivos y no deberían obtener ninguna recompensa salarial. Bajo este enfoque, un trabajador sobreeducado ganaría menos que otro similar que ocupase un puesto acorde con su formación, de manera que  $\gamma_1 < 0$  y  $\gamma_2 < 0$ . Paralelamente, un trabajador infraeducado obtendría mayores ingresos que un trabajador con el mismo nivel educativo y un puesto adecuado a su formación, siendo  $\gamma_3 > 0$  (10). Finalmente, la ecuación [3] amplía la anterior para tener en cuenta otras características del trabajador que pueden explicar los salarios percibidos.

La mayoría de las investigaciones que han analizado el efecto del desajuste educativo de los ocupados sobre los ingresos salariales percibidos han estimado por MCO alguna de las ecuaciones anteriores con datos de corte transversal. Sin embargo, con información longitudinal, el uso de métodos de estimación para datos de panel nos va a permitir controlar la denominada heterogeneidad individual inobservada, algo necesario en este tipo de análisis, tal y como se apuntaba en la introducción. La idea general que hay detrás de este planteamiento es que pueden existir características específicas individuales que, por un parte, son difíciles o incluso imposibles de observar o medir y, por otra parte, varían a través de los individuos pero no en el tiempo. Uno de los ejemplos más citados de tales características inobservadas es la inteligencia o habilidad (Griliches, 1977), de manera que parece bastante plausible que éstas se ajusten a ambos supuestos: varían a través de los individuos y, presumiblemente, permanecen más o menos constantes en el tiempo. En este contexto, los resultados por MCO podrían resultar sesgados por la omisión de variables relevantes en la ecuación, de manera que la posibilidad de tener en cuenta estas características personales es posiblemente la mayor ventaja del uso de datos de panel (Baltagi, 2001).

---

(10) Una especificación alternativa para analizar el efecto del desajuste educativo en los salarios (Duncan y Hoffman, 1981) consiste en descomponer los años de escolarización ( $E_{it}$ ), en años requeridos ( $E_{it}^r$ ), de sobreeducación ( $E_{it}^s$ ) y de infraeducación ( $E_{it}^i$ ), definiéndose la siguiente función de salarios:  $\ln Y_{it} = \beta_0 + \beta_1 E_{it}^r + \beta_2 E_{it}^s + \beta_3 E_{it}^i + \beta_4 X_{it} + u_{it}$ .

Bajo este enfoque de panel de datos, la ecuación [3] anterior puede expresarse como:

$$\ln Y_{it} = \alpha + \beta Z_{it} + u_{it} \quad [4]$$

donde  $Z_{it}$  representa al conjunto de variables explicativas del modelo y  $u_{it}$  al término de perturbación, que, se supone, está formado por los dos componentes siguientes:

$$u_{it} = \alpha_i + \eta_{it} \quad [5]$$

correspondiéndose la segunda parte,  $\eta_{it}$ , con el término de error aleatorio habitual, homoscedástico y no correlacionado ni entre individuos ni en el tiempo, mientras que la primera,  $\alpha_i$ , se conoce como efecto específico individual y en contraste con el término anterior, varía a través de los individuos pero es constante en el tiempo. De esta forma, este segundo componente, es el que representa la heterogeneidad inobservable de los individuos y la hipótesis adoptada acerca de su naturaleza estadística da lugar a los modelos de efectos fijos y aleatorios. El primero de estos modelos considera que los  $\alpha_i$  son parámetros desconocidos que tienen que ser estimados junto con el resto de parámetros, mientras que en el modelo de efectos aleatorios, los  $\alpha_i$  son tratados como extracciones de una distribución de probabilidad determinada, supuesto que resulta más realista y, por esta razón, más utilizado en los análisis empíricos(11).

En el siguiente apartado estimaremos distintas especificaciones de las ecuaciones salariales bajo este enfoque de panel de datos, utilizando un modelo de efectos aleatorios(12). Además, compararemos estos resultados con los obtenidos al considerar las siete olas del PHOGUE como una única muestra de corte transversal, es decir, sin explotar la estructura longitudinal de la base de datos.

---

(11) A cambio, el modelo de efectos aleatorios supone que éstos son independientes del resto de regresores del modelo,  $Z_{it}$ , supuesto difícil de mantener en la mayoría de los casos. Hausman y Taylor (1981) desarrollan un modelo en el que algunas de las variables explicativas están relacionadas con los efectos individuales mientras que otras no. Es un estimador instrumental cuya aplicación requiere la catalogación de las variables en exógenas o endógenas, lo que hace difícil su aplicación práctica. Por su parte, el modelo de efectos fijos no permite estimar los coeficientes de variables que permanecen constantes en el tiempo como son, entre otras, el sexo o la educación en nuestro caso, por lo que tampoco suele utilizarse en este tipo de análisis.

(12) Estos modelos se han estimado por Mínimos Cuadrados Generalizados Factibles (Greene, 1998).



#### 4. RESULTADOS DEL ANÁLISIS ECONÓMÉRICO

Las tablas 6 y 7 presentan los resultados obtenidos al estimar por MCO y con el estimador de efectos aleatorios las ecuaciones [1], [2] y [3] para varones y mujeres, respectivamente.

**Tabla 6**  
ESTIMACIÓN DE LA FUNCIÓN DE SALARIOS (1995-2001). VARONES  
(Continúa)

VARIABLES EXPLICATIVAS	MCO (POOLED)		
	(1)	(2)	(3)
Años de educación	0,079 (0,001)	0,075 (0,001)	0,034 (0,001)
Años de experiencia	0,036 (0,001)	0,035 (0,001)	0,017 (0,001)
Años experiencia <sup>2</sup> /100	-0,039 (0,002)	-0,039 (0,002)	-0,020 (0,002)
Tiene Pareja			0,070 (0,008)
Sobreeducado tipo 1		-0,060 (0,010)	-0,037 (0,008)
Sobreeducado tipo 2		-0,158 (0,011)	-0,072 (0,009)
Infraeducado		-0,092 (0,011)	-0,040 (0,009)
Antigüedad [1-2 años]			0,022 (0,011)
Antigüedad [2-5 años]			0,045 (0,011)
Antigüedad +5 años			0,097 (0,011)
Directivos			0,587 (0,020)
Profesionales y téc.superiores			0,382 (0,016)
Profesionales y téc.de apoyo			0,207 (0,014)
Empleados administrativos			0,136 (0,015)

Nota: errores estándar entre paréntesis. La variable dependiente es el logaritmo natural del salario hora bruto. Las estimaciones incluyen un término constante y en la especificación (3) se controla por estado de salud, migración, sector de actividad y región de residencia. Para todas las especificaciones, los valores del estadístico del test de Breusch Pagan-LM permiten rechazar la hipótesis de ausencia de efectos aleatorios.

**Tabla 6**  
**ESTIMACIÓN DE LA FUNCIÓN DE SALARIOS (1995-2001). VARONES**  
 (Continuación)

VARIABLES EXPLICATIVAS	MCO (POOLED)		
	(1)	(2)	(3)
Restauración			0,025 (0,013)
Industria, construcción y minería			0,076 (0,010)
Operadores			0,043 (0,012)
Trabajadores no clasificados			0,107 (0,035)
Empresa privada < 20 trabajadores			-0,218 (0,009)
Empresa privada 20-99 trabajadores			-0,116 (0,010)
Empresa privada + 100 trabajadores			0,043 (0,010)
Empresa privada y missing en nº de trabajadores			-0,204 (0,040)
Jornada a tiempo parcial			0,129 (0,022)
Contrato indefinido			0,118 (0,009)
Formación en la empresa			0,051 (0,012)
3 veces o menos parado en últimos 5 años			-0,046 (0,008)
4-6 veces parado en últimos 5 años			-0,013 (0,015)
> 6 veces parado en últimos 5 años			-0,052 (0,028)
Parado en últimos 5 años, sin determinar cuantas veces			-0,112 (0,034)
Número de observaciones	11.523	11.523	11.523

Nota: errores estándar entre paréntesis. La variable dependiente es el logaritmo natural del salario hora bruto. Las estimaciones incluyen un término constante y en la especificación (3) se controla por estado de salud, migración, sector de actividad y región de residencia. Para todas las especificaciones, los valores del estadístico del test de Breusch Pagan-LM permiten rechazar la hipótesis de ausencia de efectos aleatorios.

Tabla 6

## ESTIMACIÓN DE LA FUNCIÓN DE SALARIOS (1995-2001). VARONES

(Continuación)

VARIABLES EXPLICATIVAS	EFECTOS ALEATORIOS		
	(1)	(2)	(3)
Años de educación	0,075 (0,001)	0,074 (0,001)	0,048 (0,001)
Años de experiencia	0,037 (0,001)	0,037 (0,001)	0,024 (0,001)
Años experiencia <sup>2</sup> /100	-0,041 (0,003)	-0,041 (0,003)	-0,028 (0,002)
Tiene Pareja			0,061 (0,010)
Sobreeducado tipo 1		-0,019 (0,007)	-0,018 (0,007)
Sobreeducado tipo 2		-0,045 (0,008)	-0,037 (0,008)
Infraeducado		-0,019 (0,007)	-0,016 (0,007)
Antigüedad [1-2 años)			0,032 (0,008)
Antigüedad [2-5 años)			0,065 (0,009)
Antigüedad +5 años			0,078 (0,011)
Directivos			0,284 (0,020)
Profesionales y técnicos superiores			0,221 (0,018)
Profesionales y técnicos de apoyo			0,128 (0,014)
Empleados administrativos			0,085 (0,015)
Restauración			0,028 (0,013)
Industria, construcción y minería			0,058 (0,010)
Operadores			0,044 (0,011)
Trabajadores no clasificados			0,113 (0,042)

Nota: errores estándar entre paréntesis. La variable dependiente es el logaritmo natural del salario hora bruto. Las estimaciones incluyen un término constante y en la especificación (3) se controla por estado de salud, migración, sector de actividad y región de residencia. Para todas las especificaciones, los valores del estadístico del test de Breusch Pagan-LM permiten rechazar la hipótesis de ausencia de efectos aleatorios.

**Tabla 6**  
**ESTIMACIÓN DE LA FUNCIÓN DE SALARIOS (1995-2001). VARONES**  
 (Conclusión)

<i>VARIABLES EXPLICATIVAS</i>	<i>EFFECTOS ALEATORIOS</i>		
	<i>(1)</i>	<i>(2)</i>	<i>(3)</i>
Empresa privada < 20 trabajadores			-0,130 (0,011)
Empresa privada 20-99 trabajadores			-0,062 (0,012)
Empresa privada + 100 trabajadores			0,003 (0,012)
Empresa privada y missing en nº trabajadores			-0,152 (0,034)
Jornada a tiempo parcial			0,189 (0,020)
Contrato indefinido			0,082 (0,007)
Formación en la empresa			0,027 (0,009)
3 veces o menos parado en últimos 5 años			-0,076 (0,011)
4-6 veces parado en últimos 5 años			-0,046 (0,022)
> 6 veces parado en últimos 5 años			-0,103 (0,040)
Parado en últimos 5 años, sin determinar cuántas veces			-0,153 (0,046)
Número de observaciones	11.523	11.523	11.523

Nota: errores estándar entre paréntesis. La variable dependiente es el logaritmo natural del salario hora bruto. Las estimaciones incluyen un término constante y en la especificación (3) se controla por estado de salud, migración, sector de actividad y región de residencia. Para todas las especificaciones, los valores del estadístico del test de Breusch Pagan-LM permiten rechazar la hipótesis de ausencia de efectos aleatorios.

**Tabla 7**  
ESTIMACIÓN DE LA FUNCIÓN DE SALARIOS (1995-2001). MUJERES  
(Continúa)

VARIABLES EXPLICATIVAS	MCO (POOLED)		
	(1)	(2)	(3)
Años de educación	0,089 (0,001)	0,082 (0,001)	0,032 (0,001)
Años de experiencia	0,039 (0,001)	0,036 (0,001)	0,015 (0,001)
Años experiencia <sup>2</sup> /100	-0,053 (0,003)	-0,051 (0,003)	-0,022 (0,003)
Tiene Pareja			0,037 (0,008)
Sobreeducado tipo 1		-0,129 (0,012)	-0,045 (0,010)
Sobreeducado tipo 2		-0,268 (0,014)	-0,086 (0,013)
Infraeducado		-0,185 (0,016)	-0,058 (0,014)
Antigüedad [1-2 años)			0,016 (0,014)
Antigüedad [2-5 años)			0,055 (0,014)
Antigüedad +5 años			0,148 (0,014)
Directivos			0,553 (0,038)
Profesionales y técnicos superiores			0,047 (0,020)
Profesionales y técnicos de apoyo			0,263 (0,018)
Empleados administrativos			0,146 (0,016)
Restauración			0,052 (0,013)
Industria, construcción y minería			0,050 (0,021)
Operadores			0,049 (0,026)
Trabajadores no clasificados			-0,025 (0,110)

Nota: errores estándar entre paréntesis. La variable dependiente es el logaritmo natural del salario hora bruto. Las estimaciones incluyen un término constante y en la especificación (3) se controla por estado de salud, migración, sector de actividad y región de residencia. Para todas las especificaciones, los valores del estadístico del test de Breusch Pagan-LM permiten rechazar la hipótesis de ausencia de efectos aleatorios.

**Tabla 7**  
**ESTIMACIÓN DE LA FUNCIÓN DE SALARIOS (1995-2001). MUJERES**  
 (Continuación)

<i>VARIABLES EXPLICATIVAS</i>	<i>MCO (POOLED)</i>		
	<i>(1)</i>	<i>(2)</i>	<i>(3)</i>
Empresa privada < 20 trabajadores			-0,291 (0,011)
Empresa privada 20-99 trabajadores			-0,136 (0,013)
Empresa privada + 100 trabajadores			-0,060 (0,014)
Empresa privada y missing en nº trabajadores			-0,224 (0,053)
Jornada a tiempo parcial			0,119 (0,011)
Contrato indefinido			0,139 (0,011)
Formación en la empresa			0,015 (0,013)
3 veces o menos parado en últimos 5 años			-0,043 (0,001)
4-6 veces parado en últimos 5 años			-0,090 (0,022)
> 6 veces parado en últimos 5 años			-0,084 (0,044)
Parado en últimos 5 años, sin determinar cuantas veces			-0,041 (0,040)
Número de observaciones	6.700	6.700	6.700

Nota: errores estándar entre paréntesis. La variable dependiente es el logaritmo natural del salario hora bruto. Las estimaciones incluyen un término constante y en la especificación (3) se controla por estado de salud, migración, sector de actividad y región de residencia. Para todas las especificaciones, los valores del estadístico del test de Breusch Pagan-LM permiten rechazar la hipótesis de ausencia de efectos aleatorios.

Tabla 7

ESTIMACIÓN DE LA FUNCIÓN DE SALARIOS (1995-2001). MUJERES  
(Continuación)

VARIABLES EXPLICATIVAS	EFECTOS ALEATORIOS		
	(1)	(2)	(3)
Años de educación	0,085 (0,002)	0,083 (0,002)	0,047 (0,002)
Años de experiencia	0,038 (0,001)	0,0382 (0,001)	0,020 (0,001)
Años experiencia <sup>2</sup> /100	-0,054 (0,004)	-0,054 (0,004)	-0,031 (0,004)
Tiene Pareja			0,045 (0,011)
Sobreeducado tipo 1		-0,031 (0,009)	-0,020 (0,008)
Sobreeducado tipo 2		-0,076 (0,011)	-0,049 (0,011)
Infraeducado		-0,040 (0,012)	-0,025 (0,011)
Antigüedad [1-2 años)			0,026 (0,010)
Antigüedad [2-5 años)			0,080 (0,012)
Antigüedad +5 años			0,132 (0,014)
Directivos			0,337 (0,037)
Profesionales y técnicos superiores			0,319 (0,022)
Profesionales y técnicos de apoyo			0,147 (0,019)
Empleados administrativos			0,105 (0,017)
Restauración			0,015 (0,015)
Industria, construcción y minería			0,041 (0,021)
Operadores			0,046 (0,024)
Trabajadores no clasificados			0,067 (0,118)

Nota: errores estándar entre paréntesis. La variable dependiente es el logaritmo natural del salario hora bruto. Las estimaciones incluyen un término constante y en la especificación (3) se controla por estado de salud, migración, sector de actividad y región de residencia. Para todas las especificaciones, los valores del estadístico del test de Breusch Pagan-LM permiten rechazar la hipótesis de ausencia de efectos aleatorios.

**Tabla 7**  
ESTIMACIÓN DE LA FUNCIÓN DE SALARIOS (1995-2001). MUJERES  
(Conclusión)

<i>VARIABLES EXPLICATIVAS</i>	<i>EFFECTOS ALEATORIOS</i>		
	(1)	(2)	(3)
Empresa privada < 20 trabajadores			-0,215 (0,013)
Empresa privada 20-99 trabajadores			-0,106 (0,014)
Empresa privada + 100 trabajadores			-0,074 (0,015)
Empresa privada y missing en nº trabajadores			-0,242 (0,046)
Jornada a tiempo parcial			0,155 (0,011)
Contrato indefinido			0,094 (0,009)
Formación en la empresa			0,008 (0,010)
3 veces o menos parado en últimos 5 años			-0,073 (0,013)
4-6 veces parado en últimos 5 años			-0,144 (0,030)
> 6 veces parado en últimos 5 años			-0,053 (0,060)
Parado en últimos 5 años, sin determinar cuántas veces			-0,121 (0,061)
Número de observaciones	6.700	6.700	6.700

Nota: errores estándar entre paréntesis. La variable dependiente es el logaritmo natural del salario hora bruto. Las estimaciones incluyen un término constante y en la especificación (3) se controla por estado de salud, migración, sector de actividad y región de residencia. Para todas las especificaciones, los valores del estadístico del test de Breusch Pagan-LM permiten rechazar la hipótesis de ausencia de efectos aleatorios.

Las tres primeras columnas de dichas tablas recogen las estimaciones mínimo-cuadráticas correspondientes a las siete olas consideradas conjuntamente como una única muestra de corte transversal (pooled). Los principales resultados derivados de esta estrategia de estimación se discuten a continuación. De acuerdo con la columna (1), el rendimiento medio de la educación en el periodo 1995-2001 se sitúa en el 7,9% para los varones y en el 8,9% para las mujeres, resultado consistente



con los obtenidos para el mercado de trabajo español en otros estudios referidos aproximadamente al mismo periodo temporal (Barceinas-Paredes *et al.*, 2001 y Telhado y Silva, 2002), y ligeramente por debajo de los estimados para el comienzo de la década de los ochenta (Lassibille, 1993), en línea con la tendencia de descenso registrada en otros países (Trostel *et al.*, 2002).

Según la columna (2), los trabajadores que presentan desajuste educativo perciben, por término medio, un salario inferior al de aquellos que poseen un nivel de estudios acorde con las exigencias del puesto que ocupan. El coeficiente negativo obtenido para la sobreeducación está en línea con la literatura previa y apoyaría la predicción de la teoría de la competencia por los puestos de trabajo (Thurrow, 1975). Sin embargo, no resulta habitual encontrar evidencia empírica de que los trabajadores infraeducados registren también una penalización salarial frente a los adecuadamente educados con igual nivel educativo. Así, Alba (1993), al estimar una ecuación comparable a la nuestra con datos procedentes de la ECVT (1985), comprueba que los sobreeducados perciben un 17% menos que los adecuadamente educados y los infraeducados un 12,5% más respecto al mismo grupo<sup>(13)</sup>. Sin embargo, Alba y Blázquez (2004), obtienen resultados similares a los nuestros, con la misma fuente de datos y una muestra de corte transversal correspondiente al año 1995.

Puede comprobarse además que dentro del grupo de sobreeducados se detectan importantes diferencias, siendo los de tipo 2 los que sufren una mayor penalización salarial, lo cual vuelve a poner de manifiesto la variabilidad existente dentro de este colectivo y el interés de haber diferenciado dos tipos de trabajadores sobreeducados. Por otra parte, se observa que la reducción salarial por desajuste educativo es superior en el caso de las mujeres.

La columna (3) muestra que al controlar por la heterogeneidad observada en la muestra, incluyendo regresores adicionales, el rendimiento de la educación disminuye de forma importante, en línea con los resultados obtenidos en ecuaciones comparables (Alba, 1993; Trostel *et al.*, 2002). Además, las diferencias salariales entre los colectivos con y sin desajuste educativo se reducen significativamente. Así, si de acuerdo con la columna (2) entre los varones la penalización salarial era del 6%, 15,8% y 9,2% para los sobreeducados tipo 1, tipo 2 e infraeducados,

---

(13) Por su parte, Beneito *et al.* (1996) con una muestra de la ECBC (1991) y un modelo de regresión *switching* endógeno para diferenciar los regímenes salariales de los trabajadores que han participado alguna vez en los cursos de formación específica de los que no lo han hecho, estiman la formulación propuesta por Duncan y Hoffman (1981), por lo que sus resultados no son comparables a los nuestros. En cualquier caso, confirman los hechos estilizados obtenidos en la literatura empírica en este tipo de especificación, es decir, rendimiento marginal positivo para los años de sobreeducación pero menor que el de los años requeridos, y negativo para la infraeducación.

respectivamente, al incluir otros factores explicativos de los ingresos (columna 3), estos porcentajes caen en torno al 50% y pasan a ser del 3,7%, 7,2% y 4% para cada uno de estos colectivos. Como puede observarse, estas reducciones son aún más importantes en el caso de las mujeres.

Las tres últimas columnas de las tablas 6 y 7 recogen los resultados obtenidos al estimar las tres especificaciones consideradas bajo el modelo de efectos aleatorios. El hecho más destacable a subrayar es que la magnitud de los coeficientes obtenidos para las variables ficticias que recogen el desajuste educativo se reducen aproximadamente a la mitad frente a la estimación por MCO, en cualquiera de las especificaciones consideradas y para ambos sexos, disminuyendo además su significatividad. Así, tomando como referencia la ecuación de ingresos más amplia, la correspondiente a la especificación [3], puede observarse que en el caso de los varones un trabajador sobreeducado tipo 1 percibe un 1,8% menos que otro con las mismas características que ocupe un puesto acorde con su formación. Para los infraeducados se obtiene un porcentaje muy similar, con un 1,6% de penalización salarial, mientras que ésta se duplica para los sobreeducados tipo 2, situándose en el 3,7%. Los coeficientes en el colectivo femenino también presentan esta pauta de reducción, aunque las penalizaciones salariales siguen siendo mayores en este grupo, con porcentajes del 2 y 4,9% para las sobreeducadas (tipo 1 y tipo 2, respectivamente) y del 2,5% para las infraeducadas.

Otro resultado destacable es que si comparamos la magnitud de estos coeficientes con los obtenidos para el resto de variables explicativas incluidas en la ecuación, podemos ver que las características del emparejamiento trabajador-empresa (jornada laboral, tipo de contrato, categoría ocupacional, tamaño de empresa, etc.) son los factores más determinantes a la hora de explicar la variabilidad salarial entre los trabajadores, situándose el desajuste educativo del trabajador en una posición secundaria en cuanto a la capacidad explicativa de dicha variabilidad.

Finalmente, y con objeto de contrastar si las conclusiones anteriores se mantienen para el colectivo juvenil, el más afectado por el fenómeno de la sobreeducación (Aguilar, 2005), se ha vuelto a estimar la ecuación [3] bajo la estructura de efectos aleatorios distinguiendo dos cohortes de edad: por un lado, trabajadores con 35 años como máximo y, por otro, el resto de ocupados. Los resultados obtenidos, que se recogen en la tabla 8, muestran una mayor penalización por desajuste educativo entre los trabajadores jóvenes, como cabía esperar, aunque sólo resulta significativa estadísticamente en el caso de los sobreeducados tipo 2.

De esta forma, estos resultados vuelven a apoyar la hipótesis de que los coeficientes asociados al desajuste educativo pueden estar recogiendo la heterogeneidad inobservada de los individuos y no sólo la penalización monetaria asociada a tal desajuste, siendo, por tanto, son consistentes con los obtenidos por Bauer (2002) para el mercado de trabajo alemán.

**Tabla 8**  
ESTIMACIÓN DE LA FUNCIÓN DE SALARIOS POR EDAD Y SEXO.  
EFECTOS ALEATORIOS.(1995-2001).

	$\leq 35$ AÑOS		$>35$ AÑOS	
	VARONES	MUJERES	VARONES	MUJERES
Educación (años)	0,031 (0,002)	0,041 (0,002)	0,060 (0,002)	0,058 (0,003)
Experiencia (años)	0,035 (0,003)	0,022 (0,004)	0,020 (0,003)	0,012 (0,00)
Experiencia <sup>2</sup> /100 (años <sup>2</sup> )	-0,103 (0,015)	-0,052 (0,021)	-0,017 (0,005)	-0,016 (0,006)
Sobreeducado tipo 1	-0,020 (0,013)	-0,026 (0,013)	-0,015 (0,008)	-0,012 (0,011)
Sobreeducado tipo 2	-0,051 (0,013)	-0,060 (0,016)	-0,027 (0,010)	-0,037 (0,015)
Infraeducado	-0,015 (0,014)	-0,030 (0,018)	-0,019 (0,009)	-0,018 (0,014)
+ Regresores adicionales				
Número de observaciones	5.049	3.544	6.474	3.156

Nota: errores estándar entre paréntesis. Las estimaciones incluyen todos los regresores de la especificación (3) de las tablas 6 y 7. Los valores del estadístico del test de Breusch Pagan-LM permiten rechazar la hipótesis de ausencia de efectos aleatorios.

## CONCLUSIONES

En este trabajo se exploran las posibilidades que ofrece el PHOGUE para analizar los efectos salariales del desajuste educativo en el mercado de trabajo español. Para ello se propone una medida subjetiva del desajuste educativo y se estiman distintas especificaciones de la función de ingresos formulada por Verdugo y Verdugo (1989) para el periodo 1995-2001.

El análisis realizado ha puesto de manifiesto que la medida subjetiva de desajuste educativo construida con la información del PHOGUE no es directamente comparable con las derivadas de otras encuestas como la ECVT (1985) o la ECBC (1991), a partir de las cuales se ha estudiado tradicionalmente el desajuste educativo de los trabajadores españoles y sus efectos sobre los salarios. Así, aunque el PHOGUE permite distinguir dos tipos de trabajadores sobreeducados, sólo uno de ellos se ajusta al concepto de exceso educativo utilizado en la literatura. Por el contrario, el segundo grupo puede catalogarse de “aparentemente sobreeducado” ya que parece que su desajuste no se deriva de la formación recibida sino de otras características personales, de manera que los resultados relativos a este grupo deben ser interpretados con cautela.

Las estimaciones iniciales minimocuadráticas de la ecuación de ingresos, obtenidas considerando conjuntamente los siete años analizados, indican que el desajuste educativo implica una penalización salarial en torno al 4% en el caso de los trabajadores sobreeducados en un sentido estricto (los catalogados como tipo 1) y de los infraeducados, y por encima del 7% en el caso de los trabajadores aparentemente sobreeducados. Pero el resultado más importante es que al explotar la naturaleza longitudinal del PHOGUE teniendo en cuenta los efectos individuales inobservados mediante el modelo de efectos aleatorios, los coeficientes de las variables de desajuste educativo se reducen de forma notable –en torno al 50%- e incluso dejan de ser significativos en algunos casos. Estos resultados son consistentes con los obtenidos por Bauer (2002) para el mercado de trabajo alemán e indican que es importante tener en cuenta la heterogeneidad inobservada al analizar los efectos salariales del desajuste educativo. Por ello, consideramos como una posible extensión de este trabajo la ampliación del análisis a otros países europeos con el fin de confirmar los resultados obtenidos para España, dada la disponibilidad actual de fuentes estadísticas comparables.

## REFERENCIAS

- AGUILAR, M.I. (2005). «La inserción laboral de los jóvenes. Un enfoque microeconómico». Ed. Thomson-Civitas. *Estudios y Monografías de Economía*. Navarra, 2005.
- ALBA, A. (1993). «Mismatch in the Spanish labor market. Overeducation?». *Journal of Human Resources*, Vol. 28, nº 2, pp. 259-278.
- ALBA, A. Y BLÁZQUEZ, M. (2004). «Types of job match, overeducation and labour mobility in Spain» en *Overeducation in Europe: current issues in theory and policy*, eds. Büchel, F., Grip, A. y Mertens, A. Edward Elgar, Cheltenham.
- BALTAGI, B. (2001). «Econometric Analysis of Panel Data». 2<sup>nd</sup> edition. Wiley, New York.
- BARCEINAS-PAREDES, F.; OLIVER-ALONSO, J.; RAYMOND-BARA, J.L. Y ROIG-SABATÉ, J.L. (2001). «Education and Earnings in Europe», capítulo 13, eds. Harmon, C., Walker, I. y Westergaard-Nielsen, N. Edward Elgar, USA.
- BAUER, T.K. (2002). «Educational mismatch and wages: a panel analysis». *Economics of Education Review*, vol. 21, pp. 221-229.
- BENEITO, P.; FERRI, J.; MOLTÓ, M.L. Y URIEL, E. (1996). «Desajuste educativo y formación laboral especializada: efectos sobre los rendimientos salariales». *Documento de trabajo WP-EC 9611*, Instituto Valenciano de Investigaciones Económicas
- CHEVALIER, A. (2003). «Measuring over-education». *Económica*, vol. 70, nº 3, pp. 509-531.
- DUNCAN, G. Y HOFFMAN, S. D. (1981). «The incidence and wage effects of overeducation». *Economics of Education Review*, vol. 1, nº 1, pp. 75-86.
- FERSTERER, J. Y WINTER-EBMER, R. (2003). «Are Austrian returns to education falling over time?». *Labour Economics*, 10, pp. 73-89.
- GARCÍA MONTALVO, J. (1995). «Empleo y sobrecualificación: el caso español». *Documento de Trabajo 95-20*, FEDEA.
- GARCÍA-SERRANO, C. Y MALO, M.A. (1996). «Desajuste educativo y movilidad laboral en España». *Revista de Economía Aplicada*, vol. 4, nº 11, pp. 105-131.
- GARCÍA-SERRANO, C. Y MALO, M.A. (1997). «¿Es diferente el desajuste educativo de las mujeres?». *Información Comercial Española*, nº 760, pp. 117-128.

- GAUTIER, P.A., VAN DEN BERG, G.J., VAN OURS, J.C. AND RIDDER, G. (2002). «Worker turnover at the firm level and crowding out of lower educated workers». *European Economic Review*, nº 46, pp. 523-538.
- GREENE, W. (1998). «Análisis econométrico». *Prentice-Hall Iberia*, Madrid.
- GRILICHES, Z. (1977). «Estimating the returns to schooling: some econometrics problems». *Econometrica*, 45, pp.1-22.
- GROOT, W. Y MAASSEN VAN DEN BRINK, H. (2000). «Overeducation in the labor market: a meta-analysis». *Economics of Education Review*, vol. 19, nº 2, pp. 149-158.
- HARTOG, J. (2000). «Over-education and earnings: where are we, where should we go?». *Economics of Education Review*, vol. 19, nº 2, pp. 131-147.
- HAUSMAN, J. Y TAYLOR, W. (1981). «Panel data and unobservable individual effects». *Econometrica*, 49, pp. 1377-1398.
- KIKER, B. F.; SANTOS, M. C. Y MENDES DE OLIVEIRA, M. (1997). «Overeducation and undereducation: evidence for Portugal». *Economics of Education Review*, vol. 16, nº 2, pp. 111-125.
- LASSIBILLE, G. (1993). «El rendimiento de las inversiones educativas en España». *Estadística Española*, vol. 35, nº 134, pp. 645-663.
- MUNDLAK, Y. (1978). «On the pooling of time series and cross sectional data». *Econometrica*, 46, pp. 69-86.
- PERACCHI, F. (2002). «The European Community Household Panel: a review». *Empirical Economics*, 27, pp. 63-90.
- ROBST, J. (1995). «College quality and overeducation». *Economics of Education Review*, nº 14, (3), pp. 221-228.
- TELHADO, P. Y SILVA, P. (2002). «Is there a return-risk link in education?». *Economics Letters*, 75, pp. 31-37.
- THURROW, L.C. (1975). «Generating inequality. New York.» *Basic Books*.
- TROSTEL, P.; WALKER, I. Y WOOLLEY, P. (2002). «Estimates of the economic return to schooling for 28 countries» *Labour Economics*, 9, pp. 1-16.
- VERDUGO, R. Y VERDUGO, N. (1989). «The impact of surplus schooling on earnings». *Journal of Human Resources*, vol. 24, nº 4, pp. 629-643.
- WALKER, I. Y ZHU, Y. (2005). «The college wage premium, overeducation and the expansion of higher education in the UK». IZA Discussion Paper, nº 1627.

## APÉNDICE

**Tabla A.1**  
DESCRIPCIÓN ESTADÍSTICA DE LAS VARIABLES  
(MEDIA Y DESVIACIÓN ESTÁNDAR). 1995-2001

(Continúa)

	Sobreducados Tipo 1	Sobreducados Tipo 2	Adecuadamente educados	Infraeducados
Salario hora bruto (€ 2001)*	8,81 (5,64)	5,87 (3,01)	9,33 (5,45)	5,97 (3,00)
Log (salario hora bruto)*	2,017 (0,554)	1,666 (0,446)	2,090 (0,530)	1,688 (0,436)
Varón	0,587 (0,492)	0,651 (0,476)	0,589 (0,492)	0,722 (0,447)
Edad (años)*	37,011 (9,934)	34,604 (10,088)	40,117 (10,350)	38,494 (11,441)
Jóvenes (≤35 años)	0,483 (0,499)	0,593 (0,491)	0,359 (0,480)	0,431 (0,495)
Tiene pareja	0,674 (0,468)	0,636 (0,480)	0,731 (0,443)	0,703 (0,456)
Educación (años)*	12,140 (3,943)	9,179 (3,426)	11,081 (4,281)	7,153 (2,617)
Experiencia (años)*	16,925 (11,291)	16,378 (11,702)	20,559 (12,059)	21,887 (13,069)
Antigüedad < 1 año	0,185 (0,388)	0,291 (0,454)	0,142 (0,349)	0,239 (0,426)
Antigüedad [1-2 años]	0,123 (0,328)	0,159 (0,365)	0,092 (0,290)	0,139 (0,346)
Antigüedad [2-5 años]	0,152 (0,359)	0,174 (0,379)	0,131 (0,337)	0,167 (0,373)
Antigüedad + 5 años	0,538 (0,498)	0,375 (0,484)	0,633 (0,481)	0,453 (0,497)
Directivos	0,037 (0,191)	0,010 (0,103)	0,037 (0,189)	0,013 (0,116)
Profesionales y téc. superiores	0,240 (0,427)	0,021 (0,145)	0,288 (0,453)	0,010 (0,102)
Profesionales y téc. de apoyo	0,161 (0,367)	0,078 (0,269)	0,129 (0,335)	0,044 (0,206)
Empleados administrativos	0,155 (0,362)	0,083 (0,276)	0,109 (0,311)	0,049 (0,216)
Restauración	0,123 (0,328)	0,195 (0,396)	0,117 (0,321)	0,152 (0,359)
Industria, construcción y minería	0,133 (0,340)	0,223 (0,416)	0,164 (0,370)	0,343 (0,475)
Operadores	0,062 (0,241)	0,135 (0,342)	0,077 (0,267)	0,148 (0,356)
Trabajadores no cualificados	0,076 (0,266)	0,247 (0,431)	0,068 (0,251)	0,234 (0,423)
Trabajadores no clasificados	0,009 (0,097)	0,003 (0,055)	0,008 (0,090)	0,001 (0,041)

Nota: todas las variables son binarias excepto las que aparecen con asterisco.

**Tabla A.1**  
DESCRIPCIÓN ESTADÍSTICA DE LAS VARIABLES  
(MEDIA Y DESVIACIÓN ESTÁNDAR). 1995-2001

(Continuación)

	Sobreducados Tipo 1	Sobreducados Tipo 2	Adecuadamente educados	Infraeducados
Empresa Pública	0,339 (0,473)	0,138 (0,345)	0,356 (0,478)	0,101 (0,301)
Privada <20	0,318 (0,465)	0,475 (0,499)	0,289 (0,453)	0,496 (0,500)
Privada 20-99	0,169 (0,374)	0,221 (0,415)	0,186 (0,389)	0,235 (0,424)
Privada +100	0,169 (0,375)	0,158 (0,365)	0,162 (0,369)	0,157 (0,363)
Privada missing en tamaño	0,004 (0,065)	0,005 (0,076)	0,005 (0,071)	0,009 (0,097)
Jornada parcial	0,058 (0,235)	0,093 (0,291)	0,053 (0,224)	0,061 (0,240)
Contrato indefinido	0,707 (0,454)	0,535 (0,498)	0,759 (0,427)	0,567 (0,495)
Industria	0,272 (0,445)	0,412 (0,492)	0,288 (0,453)	0,524 (0,499)
Servicios	0,727 (0,445)	0,587 (0,492)	0,711 (0,453)	0,475 (0,499)
Formación en la empresa	0,123 (0,329)	0,056 (0,231)	0,106 (0,308)	0,024 (0,154)
Sin paro durante los últimos 5 años	0,661 (0,473)	0,514 (0,499)	0,716 (0,450)	0,546 (0,497)
3 veces o menos parado en los últimos 5 años	0,286 (0,451)	0,408 (0,491)	0,237 (0,425)	0,363 (0,481)
4-6 veces parado en los últimos 5 años	0,035 (0,185)	0,056 (0,230)	0,032 (0,177)	0,059 (0,236)
> 6 veces parado en los últimos 5 años	0,008 (0,090)	0,011 (0,108)	0,007 (0,088)	0,018 (0,135)
Parado en los últimos 5 años, sin determinar cuantas veces	0,008 (0,008)	0,009 (0,097)	0,005 (0,076)	0,011 (0,107)

Nota: todas las variables son binarias excepto las que aparecen con asterisco.



**Tabla A.1**  
DESCRIPCIÓN ESTADÍSTICA DE LAS VARIABLES  
(MEDIA Y DESVIACIÓN ESTÁNDAR). 1995-2001

(Conclusión)

	Sobreducados Tipo 1	Sobreducados Tipo 2	Adecuadamente educados	Infraeducados
Sin cambiar la CCAA de residencia	0,713 (0,452)	0,765 (0,423)	0,701 (0,457)	0,753 (0,431)
Buen estado de salud	0,985 (0,119)	0,9758 (0,153)	0,973 (0,160)	0,963 (0,187)
Región norte	0,134 (0,340)	0,117 (0,321)	0,134 (0,341)	0,107 (0,310)
Región noreste	0,146 (0,353)	0,149 (0,356)	0,189 (0,392)	0,182 (0,386)
Región Madrid	0,157 (0,364)	0,114 (0,318)	0,120 (0,325)	0,078 (0,269)
Región centro	0,116 (0,320)	0,142 (0,349)	0,122 (0,327)	0,138 (0,345)
Región este	0,241 (0,428)	0,241 (0,427)	0,222 (0,416)	0,212 (0,408)
Región sur	0,134 (0,341)	0,172 (0,378)	0,132 (0,339)	0,212 (0,409)
Región canarias	0,063 (0,243)	0,052 (0,223)	0,071 (0,257)	0,061 (0,239)
No responde región	0,006 (0,080)	0,008 (0,092)	0,006 (0,081)	0,006 (0,087)
Número de observaciones	6.088	4.272	3.896	3.967

Nota: todas las variables son binarias excepto las que aparecen con asterisco.

Fuente: elaboración propia a partir del PHOGUE (1995-2001).

## EDUCATIONAL MISMATCH AND WAGES IN SPAIN: NEW EVIDENCE WITH PANEL DATA

### ABSTRACT

This paper measures the educational mismatch and the wage effects of the overeducation and undereducation in Spain. By using the ECHP (1995-2001), pooled OLS estimated earning functions confirm the main conclusions of the existing empirical literature: the wage penalty of educational mismatch. Overeducated and undereducated workers earn less than workers with the same level of education who work in occupations for which they are adequately educated. Nevertheless, this penalty decreases significantly when we control for unobserved individual characteristics by the random effects estimator. Therefore, we conclude that taking account for unobserved heterogeneity is important when analyzing the wage effects of educational mismatch because the results may differ compared with those based on cross-section data.

*Key words:* Educational mismatch, return to education, panel data.

*AMS Classification:* 62P20, 91B40.