

## **Sobreeducación y mercados de trabajo locales en España**

Esteban Sanromá<sup>1</sup>, Raúl Ramos<sup>2</sup>

*Universidad de Barcelona*

### **ABSTRACT**

Un aspecto relativamente poco tratado por la literatura relativa a la sobreeducación es la relación entre ésta y el territorio. La vinculación entre ambas se halla en la hipótesis de la sobreeducación diferencial, es decir, la existencia de diferencias entre hombres y mujeres en términos de sobreeducación como resultado de las restricciones geográficas de las mujeres casadas a la hora de buscar trabajo. En este contexto, Büchel y Van Ham (2002) han puesto de manifiesto la importancia de la movilidad espacial en los mercados de trabajo regionales (básicamente, el commuting) como factor explicativo de la sobreeducación diferencial femenina, pero también de la sobreeducación a nivel global.

El objetivo de este trabajo consiste en contrastar la influencia de variables individuales y de algunas características de los mercados de trabajo regionales relacionadas con la movilidad espacial de la mano de obra sobre el fenómeno de la sobreeducación en el caso concreto de la economía española a partir de los datos individuales de la Encuesta de Presupuesto Familiares de 1990-91. Los resultados no avalan la hipótesis de la sobreeducación diferencial, pero confirman que tanto el tamaño del mercado de trabajo local como la posibilidad de ampliar el ámbito de búsqueda a través del commuting son factores importantes a la hora de explicar la sobreeducación en el mercado de trabajo español.

---

<sup>1</sup> Institut d'Economia de Barcelona.

<sup>2</sup> Grup d'Anàlisi Quantitativa Regional.

## **1. INTRODUCCIÓN Y OBJETIVOS**

Un aspecto relativamente poco tratado por la literatura relativa a la sobreeducación es la relación entre ésta y el territorio. La vinculación entre ambas se halla en la hipótesis de la sobreeducación diferencial. La idea es que la sobreeducación es un fenómeno que afectaría básicamente a las mujeres casadas, dado que su proceso de búsqueda de empleo queda restringido al mercado local donde reside, a diferencia del marido que puede buscar un puesto de trabajo más adecuado a su formación en un mercado más amplio.

Algunos estudios recientes, como Büchel y Battu (2002) o Büchel y Van Ham (2002) han puesto de manifiesto la importancia de los mercados de trabajo regionales (básicamente aspectos relacionados con el commuting) como factor a tener en cuenta en la explicación de la sobreeducación y no sólo de la sobreeducación diferencial de las mujeres casadas. De hecho, Büchel y Van Ham (2002) obtienen evidencia empírica sobre el hecho de que variables de carácter regional relacionadas con la distribución espacial de los empleos y con el tamaño del mercado de trabajo local afectan a la sobreeducación. Es más, la posibilidad de acceder a zonas geográficas más amplias disminuyen la probabilidad de estar sobreeducado para todos los individuos (y no sólo para las mujeres casadas).

El objetivo de este trabajo consiste en contrastar la influencia de variables individuales y de algunas características de los mercados de trabajo regionales relacionadas con la movilidad espacial de la mano de obra sobre el fenómeno de la sobreeducación en el caso concreto de la economía española. Para ello se utilizan datos individuales de la Encuesta de Presupuesto Familiares de 1990-91. Los resultados no avalan la hipótesis de la sobreeducación diferencial, pero confirman que tanto el tamaño del mercado de trabajo local como la posibilidad de ampliar el ámbito de búsqueda a través del commuting son factores importantes a la hora de explicar la sobreeducación en el mercado de trabajo español

El resto del trabajo se estructura de la siguiente manera. En primer lugar, se realiza una breve revisión de la literatura. En el tercer apartado se describe la base de datos y, a partir de ella, se calcula la incidencia de la sobreeducación en el mercado de trabajo español. En el cuarto se comenta la metodología econométrica utilizada y se presentan los principales resultados. Por último, se recogen las principales conclusiones del trabajo y las líneas de avance del mismo.

## **2. BREVE REVISIÓN DE LA LITERATURA**

A partir de las bases fijadas por Richard Freeman (1976) y por Duncan y Hoffman (1981), el fenómeno de la sobreeducación ha sido analizado en diversos países desarrollados<sup>3</sup>. El caso español no ha sido distinto y desde el trabajo de Alba (1993), diversos autores han analizado la importancia de este fenómeno para la economía española. Sin embargo, un aspecto que ha merecido escasa atención en la abundante literatura relativa a la sobreeducación es la relación de esta variable con el territorio. La única consideración de los aspectos espaciales de la sobreeducación se encuentra en la hipótesis denominada de la sobreeducación diferencial. El trabajo pionero en este ámbito es el publicado por Frank en 1978 para la economía estadounidense. Dicho estudio tenía como objetivo explicar las diferencias salariales entre hombres y mujeres a partir del fenómeno de la sobreeducación. Frank (1978) supone que los individuos buscan trabajo tratando de maximizar sus ingresos, para lo cual minimizan la sobreeducación, es decir, la diferencia entre sus estudios y los requerimientos educativos de la vacante. Este proceso de búsqueda tiene lugar en el mercado global para los individuos solteros -dado que pueden emigrar-, pero no así para las parejas. En este caso emerge una decisión conjunta consistente en que el miembro de mayor nivel educativo y que trabaja un mayor número de horas, es decir, el que puede aportar mayores ingresos a la unidad familiar (normalmente el marido), busca empleo en el mercado de trabajo global minimizando su sobreeducación. En cambio, el otro miembro de la pareja (normalmente la esposa) busca el mejor empleo posible limitada al mercado de trabajo local en que se ha establecido el matrimonio, que es el que minimiza la sobreeducación del marido. Sólo por fortuna podrá la esposa minimizar también su sobreeducación gracias a un *match* óptimo. La restricción espacial provocará normalmente que su grado de sobreeducación sea mayor que el del marido. Además, en este caso, la sobreeducación que acabe padeciendo la esposa depende del tamaño del mercado de trabajo local, ya que cuanto menor sea éste, menor será el número de vacantes a las que podrá acceder la esposa. De este modo, según Frank (1978), las diferencias salariales entre sexos estarían relacionadas con el fenómeno de la sobreeducación y los mercados de trabajo locales y aquel fenómeno afectaría predominantemente a las mujeres casadas. La evidencia obtenida por dicho autor apoya esta hipótesis, ya que la diferencia salarial entre sexos en 1967 es mayor cuanto más pequeño es el tamaño de mercado de trabajo local.

Dos décadas más tarde y en un nuevo contexto de generalización de trabajos relativos a la sobreeducación, McGoldrick y Robst (1996) realizan un test de la hipótesis de Frank

---

<sup>3</sup> En concreto, Estados Unidos, Reino Unido, Holanda, España, Portugal, Grecia, Alemania, Finlandia y Hong-Kong. Un resumen de la literatura puede consultarse en Hartog (2000) y en Groot y Maassen van den Brink (2000a).

directamente en términos de sobreeducación, en lugar de estimar ecuaciones salariales. A partir de datos estadounidenses de 1985 calculan el nivel de sobreeducación por los tres métodos y encuentran que, de un lado, las mujeres casadas presentan más sobreeducación que sus maridos, pero, de otro, que el tamaño del mercado de trabajo local no ejerce ningún efecto significativo, ni por sí mismo ni de forma multiplicativa con la variable esposa. Es decir, a diferencia de Frank (1978), no encuentran evidencia de sobreeducación diferencial, ya que no hallan ninguna relación entre el tamaño del mercado local y la probabilidad de que las mujeres casadas estén sobreeducadas.

García-Serrano y Malo (1997) contrastan dicha hipótesis para el caso español a partir de la Encuesta de Estructura, Conciencia y Biografía de Clase de 1991. Los resultados obtenidos por dichos autores tampoco confirman plenamente la hipótesis de la sobreeducación diferencial. No encuentran una incidencia clara de la variable sexo sobre el desajuste educativo, aunque dicha variable sí resulta significativa en interacción con la edad o con la existencia de cohabitación. Así pues, muestran que el desajuste educativo español presenta diferencias por sexos, aunque no viene explicado por el tamaño del mercado de trabajo local. De hecho, tanto este tamaño como la interacción mujer que vive en pareja inciden negativamente sobre la probabilidad de estar infraeducado, pero no afectan a la sobreeducación.

Más recientemente, Büchel y Battu (2002) contrastan la hipótesis de la sobreeducación diferencial utilizando datos de 1995 para la economía alemana que permiten una medición subjetiva de la misma. La contrastación parece confirmar la hipótesis, ya que la probabilidad de estar sobreeducada es mayor para las mujeres casadas que viven en áreas rurales (menos de 20.000 habitantes). Ahora bien, los autores argumentan que la probabilidad de que la mujer encuentre un trabajo adecuado a sus estudios no sólo depende del tamaño del mercado de trabajo local (el argumento original de Frank) sino también de las posibilidades de acceder a través del *commuting* a otros mercados de trabajo cercanos. Al incorporar esta variable en términos de distancia de desplazamiento diario hasta el trabajo, muestran que la probabilidad de que un individuo esté sobreeducado disminuye considerablemente a medida que aumenta la distancia que recorre diariamente para desplazarse hasta su puesto de trabajo. Además, y muy importante, se desvanece la evidencia favorable a la hipótesis de la sobreeducación diferencial. Los resultados ahora evidencian que, una vez controlada la distancia “*commutada*”, el riesgo de sobreeducación es mayor para las parejas que residen en áreas rurales, para ambos miembros de la pareja y no específicamente para la mujer.

Partiendo de esta idea, Büchel y Van Ham (2002) desarrollan un marco teórico que relaciona la existencia de sobreeducación a nivel individual (tanto para hombres como para mujeres) con la disponibilidad de oportunidades de ocupación. Siguiendo a Simpson (1992), señalan que un individuo que busca trabajo en un determinado mercado de trabajo local tiene

tres opciones cuando en dicho mercado no se le ofrecen puestos de trabajo apropiados a sus características: Una primera opción es no trabajar y seguir buscando, es decir, pasar a formar parte del colectivo de desempleados. Una segunda opción es aceptar un trabajo local cuyos requerimientos sean inferiores a su nivel de educación (el resultado, por tanto, es la sobreeducación). Por último, la tercera opción consiste en aceptar un trabajo más allá de los límites del mercado de trabajo local, realizando desplazamientos diarios a una distancia superior a la deseada. Así, el aspecto central del trabajo de Büchel y van Ham (2002) consiste en analizar hasta que punto las oportunidades de empleo en los mercados de trabajos regionales (tasa de desempleo) y la movilidad personal (disponibilidad de vehículo y duración del desplazamiento al trabajo) afectan a la probabilidad de estar sobreeducado. Al igual que ocurría en el modelo de Frank (1978), las restricciones geográficas juegan un aspecto clave en la explicación, pero las diferencias fundamentales son que Büchel y Van Ham extienden el efecto a todos los trabajadores y que abren la posibilidad de ampliar el ámbito espacial de la búsqueda de empleo mediante el *commuting*

Los resultados que obtienen para el mercado de trabajo alemán ponen de manifiesto que variables de carácter regional relacionadas con la distribución espacial de los empleos explican el fenómeno de la sobreeducación. Además, una mayor movilidad de los individuos (por disponer de coche o por realizar un desplazamiento más largo) permite ampliar el tamaño efectivo del mercado de trabajo y disminuye la probabilidad de estar sobreeducado.

Tomando como punto de partida el actual estado de conocimiento sobre el tema, el presente estudio tratará de identificar los factores que inciden en el hecho de estar sobreeducado y, más concretamente, estudiar el efecto de algunos de ellos relacionados con el territorio. En una primera fase se contrastará la hipótesis de la sobreeducación diferencial de Frank (1978) para, a continuación, analizar el efecto de algunas variables territoriales, como el tamaño del mercado local o la movilidad de los individuos.

Para el caso español, la literatura empírica relativa a las decisiones de *commuting* aporta evidencia relacionada con estos aspectos que, *a priori*, conduciría a pensar en la validez de la hipótesis de sobreeducación diferencial. Así, Artís, Romaní y Suriñach (2000) estiman con datos de 1991 para Cataluña un modelo logit multinomial para estudiar la decisión de commutar a diferentes destinos, encontrando que dicha probabilidad es inferior en el caso de las esposas, especialmente aquéllas que son madres. Romaní y otros (2001) estudian las decisiones de movilidad residencial junto a las de *commuting* confirmando el mencionado efecto del estado civil y de la maternidad. Casado (2000) obtiene similares resultados con datos de la Comunidad Valenciana de 1991 con el añadido interesante de que la probabilidad de commutar se reduce a medida que aumenta el número de hijos. Estos resultados son coherentes con un patrón de reparto asimétrico de las tareas domésticas vigente en España. Podría conjeturarse, además, que

las distancias de desplazamiento son también distintas, ya que existe evidencia de que ser cabeza de familia aumenta la probabilidad, no sólo de commutar, sino de hacerlo en transporte privado (coche) (Matas, 1991 y Barrios y Martínez, 1998). Esta evidencia relativa a la movilidad cotidiana diferencial avala el interés del contraste aquí propuesto.

No obstante, como paso previo, y a fin de enmarcar el estudio, en el siguiente apartado se presenta la base de datos utilizada en el trabajo econométrico posterior, que es la Encuesta de Presupuestos Familiares 1990-91 elaborada por el INE, y se muestra el resultado de calcular la sobreeducación para la economía española utilizando datos individuales procedentes de la misma. Posteriormente, en el apartado cuarto se abordará el estudio de los factores territoriales que pueden afectar a la probabilidad de estar sobreeducado.

### **3. LA SOBREENUCACIÓN EN LA ENCUESTA DE PRESUPUESTOS FAMILIARES DE 1990/91**

#### **3.1. La Encuesta de Presupuestos Familiares 1990/91**

La fuente de información utilizada en este trabajo ha sido la Encuesta de Presupuestos Familiares realizada por el INE y referida al año 1990-1991, la cual incorpora información particularmente útil para los estudios de microeconomía laboral. En concreto, la EPF 1990-91 informa de los datos personales, de la relación con el mercado de trabajo, de los datos familiares y de la vivienda. Al contener esta abundante información individualizada con un elevado detalle territorial de la misma, la EPF resulta extraordinariamente útil para esta investigación, puesto que permite abordar de manera satisfactoria los objetivos propuestos. Además, hasta donde conocemos, no existen estudios previos que hayan utilizado esta base de datos para analizar el fenómeno de la sobreeducación.

En este trabajo se ha utilizado información correspondiente a los individuos encuestados en edad laboral (entre 16 y 65 años) sin considerar a los residentes en Ceuta y Melilla y a aquellos individuos ocupados en las fuerzas armadas o en ocupaciones no clasificadas. A partir de este conjunto de individuos se han definido dos muestras distintas. La primera muestra, formada por los 21.359 individuos que declararon estar ocupados (trabajando en el territorio nacional) y no compaginaban su trabajo con la realización de estudios, se ha utilizado para calcular la incidencia de la sobreeducación en el mercado de trabajo español. La segunda muestra, formada por los 5.492 individuos, tanto ocupados (4.889) como parados (603), que habían finalizado los niveles educativos de FP2, COU o estudios universitarios y no continuaban estudiando, ha sido utilizada para el trabajo econométrico de la sección 4.

### 3.2. La medición del desajuste educativo

La literatura ha desarrollado tres métodos distintos para medir la sobreeducación en concreto y el desajuste educativo en general<sup>4</sup>: el método objetivo<sup>5</sup>, el método subjetivo (directo e indirecto)<sup>6</sup> y el método estadístico<sup>7</sup>.

Lamentablemente, los diferentes métodos ofrecen resultados muy diversos aun cuando se apliquen a una misma base de datos (Groot y Maassen van den Brink, 2000b y Madrigal, 2002). En general, se observa que el método estadístico subestima el desajuste educativo, especialmente en su versión basada en la media (Groot y Maassen van den Brink, 2002a). Resulta evidente, no obstante, que la aplicación del método de medición viene condicionada por la información disponible.

Precisamente por ello, y teniendo en cuenta que la Encuesta de Presupuestos Familiares no contiene información subjetiva, pero informa con detalle sobre el nivel de estudios y la ocupación desempeñada, en este trabajo se ha utilizado dicho método estadístico para obtener una medida de la sobreeducación en la economía española referida al año 1990/91. La EPF contempla 14 niveles educativos, diez de los cuales son de aplicación a la muestra de mayores de 16 años, y detalla 81 ocupaciones, lo que supone un desglose a dos dígitos, inferior al óptimo, que sería una desagregación de tres dígitos de la clasificación de ocupaciones. A causa de ello, no resulta posible aplicar el criterio de la moda propuesto por Mendes de Olivera *et al.* (2000), ya que el número de categorías en que la moda no engloba el 60% del total de individuos analizados es considerable. Ello ha llevado a utilizar el criterio de la media estadística.

Los resultados de aplicar este procedimiento para los 21.359 individuos que declararon estar ocupados (trabajando en el territorio nacional) y no compaginaban su trabajo con la realización de estudios durante el período de realización de la encuesta (1990-1991) señalan que el 14,6% estaban sobreeducados. La comparación con otros resultados disponibles para aquel período (tabla 1), permite observar que el método objetivo aplicado a la EPA arroja unos resultados claramente inferiores y que el método subjetivo aplicado a la Encuesta de Estructura, Conciencia y Biografía de Clase (EECBC) ofrece un porcentaje muy superior, situándose, pues,

---

<sup>4</sup> La presentación de los diferentes métodos y de sus principales debilidades puede consultarse en Hartog (2000), Blanco (1997) y Madrigal (2002).

<sup>5</sup> Algunos trabajos que aplican este método son Rumberger (1987), Hartog y Oosterbeek (1988), Kiker y Santos (1991), Kiker *et al.* (1997) y García Montalvo (1995).

<sup>6</sup> Algunos ejemplos son Duncan y Hoffman (1981) y Sicherman (1991) para Estados Unidos, Hartog y Oosterbeek (1988) para Holanda y Alba (1993) y García Serrano y Malo (1996) para España.

<sup>7</sup> Defensores de este método en su versión media son Verdugo y Verdugo, 1989, en versión moda Kiker *et al.* (1997), mientras que Mendes de Olivera *et al.* (2000) recomiendan una moda corregida.

los porcentajes del método estadístico en un valor intermedio. Los estudios que aplican este método -con el criterio de la media- a bases de datos del año 1991 encuentran un porcentaje muy similar al nuestro en el caso de la EECBC (15,9%) y claramente menor en la EPA (8,9%). El criterio de la moda corregida de Oliver y Raymond (dirs.) (2002) arroja un valor también cercano al nuestro. Así pues, pese a las limitaciones del método, el resultado aquí obtenido con la Encuesta de Presupuestos Familiares parece suficientemente fiable para el trabajo que se pretende desarrollar, ya que se sitúa en un punto intermedio entre las estimaciones objetivas y las subjetivas y muy próximo a las obtenidas para el mismo año a partir de la EECBC y de la EPA.

Una regularidad empírica presente en la literatura es que la presencia de la sobreeducación es más elevada entre los colectivos de mayor nivel educativo, ya que están en una situación de mayor riesgo. Se ha calculado, por ello, la importancia del fenómeno de la sobreeducación en una muestra de 4.889 trabajadores ocupados que habían finalizado los niveles educativos de FP2, COU o estudios universitarios y no continuaban estudiando. La incidencia de la sobreeducación es superior en cuatro puntos al del conjunto de todos los trabajadores: un 18.7% respecto al 14.6% del total. Por sexos, el resultado también cambia, ya que ahora la sobreeducación incide más en las mujeres (19,1%) que en los varones (18,5%). Teniendo en cuenta que el desempleo es mayor entre las mujeres, podría pensarse que una mayor tasa de paro podría empujar a las mujeres a aceptar trabajos donde los requerimientos educativos sean inferiores a los que disponen. Así pues, ambos fenómenos (oportunidades de empleo y sobreeducación) podrían estar claramente interrelacionado, aspecto que se desarrolla en el siguiente apartado.

### **3. SOBREEDUCACIÓN, MERCADOS DE TRABAJO LOCALES Y MOVILIDAD**

#### **4.1. Metodología**

En este trabajo se pretende identificar los factores relacionados con el hecho de estar sobreeducado y, especialmente, estudiar el efecto de algunos de ellos relacionados con el territorio. En concreto, se pretende explicar el comportamiento de una variable dicotómica que toma valor 1 cuando el individuo considerado está sobreeducado y valor 0 en caso contrario.

Sin embargo, cuando se pretende utilizar un modelo de regresión lineal múltiple para explicar el comportamiento de una variable endógena dicotómica, sucede que en la mayoría de las ocasiones las variables explicativas no tendrían un significado por sí mismas. En estos casos, podría resultar mucho más interesante realizar un análisis que permita responder a la pregunta: ¿cuál es la probabilidad de que la variable dicotómica tome valor 1?, es decir, ¿cuál es la

probabilidad de que se obtenga evidencia a favor de que el individuo esté sobreeducado a partir de la consideración de diferentes características del individuo, del puesto de trabajo que desempeña o de las características del territorio donde reside?. Ésta es, precisamente, la característica principal de los modelos logit, los cuales, por medio de una función de probabilidad condicionada introducida dentro de la función de verosimilitud, permiten realizar una estimación de los parámetros  $\beta$  que cumplan con las condiciones requeridas para hacer interpretaciones en términos de probabilidades. Así pues, la aplicación de un modelo logit en el contexto de este trabajo permitirá obtener estimaciones del incremento en la probabilidad de que el individuo esté sobreeducado ante variaciones marginales en las variables explicativas consideradas.

Sin embargo, un aspecto adicional que hay que considerar es la presencia entre las variables explicativas de variables relacionadas con características individuales y otras relacionadas con características regionales. Ignorar la presencia de datos de niveles distintos puede tener efectos importantes sobre la inferencia realizada a partir de las estimaciones que se obtendrían a partir de las técnicas habituales. En este contexto, la solución más habitual consiste en la aplicación y estimación de modelos *multilevel* (Goldstein, 1995; Hox, 1998 o Goldstein y Rasbash, 1996). Los resultados que se muestran en el apartado siguiente se han obtenido a partir de la especificación y estimación de este tipo de modelos<sup>8</sup>.

Un último aspecto que también hay que considerar antes de proceder al análisis empírico es que para los individuos desempleados resulta imposible conocer si están o no sobreeducados, puesto que no pueden asignarse a ninguna ocupación concreta. Sin embargo, y como se ha visto anteriormente, puede suceder que ante la perspectiva de trabajar en una situación de sobreeducación el individuo decida proseguir su búsqueda, es decir, decida no aceptar el empleo y seguir en paro. En este contexto, el análisis de los factores que determinan la sobreeducación sin considerar esta posibilidad puede llevar a resultados incorrectos. Existe el problema de sesgo de selección. Para solucionar este problema se aplica el procedimiento en dos etapas propuesto por Heckman (1979). Así, la primera etapa del procedimiento de Heckman consistiría en analizar la probabilidad de estar ocupado, mientras que la segunda analizaría la probabilidad de estar sobreeducado incluyendo como una variable explicativa más la lambda de Heckman (obtenida en la primera etapa como la inversa de la ratio de Mills<sup>9</sup>). Teniendo en

---

<sup>8</sup> Los cálculos obtenidos se han realizado utilizando la macro GLIMMIX que permite utilizar el procedimiento PROC MIXED del entorno SAS v8 para estimar modelos *multilevel* donde la variable endógena se ha definido como una variable binaria (logit). Dicha macro está disponible en <http://www.sas.com/techsup/download/stat>. Para más detalles, consultar Singer (1998).

<sup>9</sup> Esta variable recoge los efectos de todas aquellas características no controladas que están relacionadas con la decisión de sobreeducación-ocupación. La introducción de dicha variable en la XII Jornadas de la Asociación de Economía de la Educación

cuenta la naturaleza binaria de ambas variables, en las dos etapas se utilizarán los modelos logit anteriormente mencionados.

En el siguiente apartado se aplica esta metodología a los datos individuales procedentes de la Encuesta de Presupuestos Familiares 1990/91.

## 4.2. Resultados

En este apartado se presentan los resultados de estimar siguiendo la metodología expuesta anteriormente cuatro especificaciones distintas orientadas a identificar los determinantes de la ocupación y de la sobreeducación en el mercado de trabajo español.

Los resultados que se muestran en las tablas 2 y 3 se corresponden a la estimación de modelos distintos donde se contrasta para la economía española la teoría de la sobreeducación diferencial y la influencia de variables relacionadas con las características de los mercados de trabajo locales sobre el fenómeno de la sobreeducación.

En la primera columna de cada una de las tablas se presentan los resultados de la primera etapa del procedimiento de Heckman, es decir los resultados de la estimación de un modelo donde la variable endógena toma un valor 1 para aquellos individuos que están ocupados y 0 en caso contrario. La segunda columna recoge los resultados de la estimación de un modelo logit donde la variable endógena toma un valor 1 para los individuos sobreeducados y 0 en caso contrario sin tener en cuenta el posible problema de sesgo de selección. Dichos resultados se presentan para poder valorar los efectos de introducir en la estimación del modelo la lambda de Heckman. En la tercera columna de cada una de las tablas se ofrecen los resultados de la estimación de la segunda etapa del procedimiento de Heckman, es decir, se ha introducido como variable explicativa la lambda de Heckman.

Se pretende, en primer lugar, contrastar la teoría de la sobreeducación diferencial para el caso español (tabla 2). Para ello, se han introducido en cada uno de los tres modelos ahora comentados diversas variables relativas a las características del individuo, pero también variables relacionadas con el mercado de trabajo local, como el tamaño del municipio o la tasa de paro<sup>10</sup>.

El nivel de estudios y la experiencia tienen efectos positivos y significativos sobre la probabilidad de estar ocupado, mientras que el hecho de ser mujer y vivir en pareja, así como ser hijo/a y convivir con los padres reducen dicha probabilidad. Las condiciones del mercado de

---

segunda ecuación corrige los posibles problemas derivados de la omisión de variables relevantes en la estimación del resto de variables explicativas.

<sup>10</sup> La información relativa a la tasa de paro provincial procede de la Encuesta de Población Activa.

trabajo local también son relevantes, dado que la tasa de paro provincial resulta significativa con un signo negativo. A partir de estos resultados también se han calculado los valores de cada individuo para la lambda de Heckman, los cuales se utilizan posteriormente como variable explicativa en el modelo de la sobreeducción.

Los resultados presentados en la segunda y en la tercera columna de la tabla 2 se corresponden al modelo explicativo de la sobreeducción. Los resultados entre ambas estimaciones no difieren en exceso, lo que queda confirmado por la no significación de la lambda de Heckman.

Dos resultados previos merecen un breve comentario. De una parte, se confirma que la probabilidad de estar sobreeducado aumenta con el nivel de estudios. Éste es un resultado usual en la literatura, tanto internacional como española. De otra, la experiencia potencial no muestra ninguna capacidad explicativa, de manera que la mayor o menor experiencia laboral no afecta a la probabilidad de estar sobreeducado. Pese a que no se trata de un test riguroso, este resultado no confirmaría la sustituibilidad entre educación y otras formas de capital humano que postula la teoría del capital humano<sup>11</sup>.

A fin de contrastar la hipótesis de la sobreeducción diferencial se han introducido algunas interacciones entre sexo, estado civil y tamaño del municipio de residencia tomando como categoría referencia, en concreto, a las mujeres casadas que viven en municipios pequeños. La hipótesis se confirmaría si las restantes categorías presentaran un menor riesgo de sobreeducción. Pese a que la mayoría de signos son negativos, los resultados señalan que únicamente es significativamente menor la probabilidad de sobreeducción para los hombres casados que residen en un municipio grande y para las mujeres solteras que residen en un municipio grande<sup>12</sup>, no existiendo diferencias significativas en las otras categorías. Estos resultados conducirían a rechazar la hipótesis de Frank (1978), coincidiendo así con los obtenidos por McGoldrick y Robst (1996) y por García Serrano y Malo (1997) y también por Büchel y Battu (2002) tras introducir la posibilidad de *commuting*.

Los resultados parecen apuntar, por el contrario, a que la sobreeducción podrían estar más relacionada con el tamaño del mercado de trabajo local (municipio) que con el sexo o el estado civil de los individuos. Ésta es, precisamente, la hipótesis de Büchel y van Ham (2002). Siguiendo a estos autores, se analiza el efecto sobre la probabilidad de sobreeducción del tamaño del mercado laboral local y de las posibilidades de acceder a un área de búsqueda de mayor dimensión. En concreto, el tamaño del mercado local se ha aproximado por el número de

---

<sup>11</sup> Contrastes más precisos de las hipótesis del capital humano pueden consultarse, entre otros, en Sicherman (1991) para Estados Unidos, Mendes de Oliveira *et al.* (2000) para Portugal y García Serrano y Malo (1996) y Beneito *et al.* (1996) para España.

<sup>12</sup> Siempre en comparación con las mujeres con pareja que habitan en municipios pequeños.

habitantes del municipio. La información disponible en la EPF se ha resumido en dos categorías: municipio de más de 50.000 habitantes o capital de provincia y resto de municipios. La posibilidad de ampliar el ámbito espacial de búsqueda a fin de lograr un mejor *match* que evite la sobreeducación se ha tratado de recoger a través de dos variables: una individual, la disponibilidad por parte del hogar de un medio de transporte privado, y otra provincial, el número de kilómetros de carretera en relación al número de vehículos<sup>13</sup>.

Los resultados (columna 1 de la tabla 3) correspondientes a la primera etapa del procedimiento de Heckman ofrecen conclusiones similares a las anteriores. Los resultados relativos a la sobreeducación (columnas 2 y 3) confirman que a mayor nivel de estudios mayor es la probabilidad de trabajar en un empleo para el que se está sobreeducado.

La lambda de Heckman resulta significativa al 10%, lo que arrojaría una cierta evidencia favorable a que en el mercado de trabajo español se da la estrategia citada anteriormente de no aceptar un empleo para el que se está sobreeducado a fin de prorrogar la búsqueda de un puesto de trabajo óptimo. Esta estrategia resultaría comprensible al existir, en el caso español, una densa red familiar de protección frente al paro y -en 1990-91- un sistema de protección por desempleo todavía generoso en relación al contexto occidental (OCDE, 1994). Con todo, la presencia de un elevado desempleo y unos períodos de paro muy dilatados actuarían como límites a dicha estrategia y, por tanto, a la significación de la lambda.

De nuevo aquí se observa la no significación de la variable experiencia, reforzando la evidencia anterior contraria a la sustituibilidad entre componentes del capital humano. Merece ser destacado que tampoco la variable sexo parece ejercer ningún efecto sobre la sobreeducación. Pese a que las medias muestrales de las diferentes encuestas o fuentes estadísticas suelen arrojar valores diferentes para ambos sexos, se confirma una vez más que, una vez controlados los restantes factores explicativos, no existen diferencias significativas<sup>14</sup>.

En el caso de las variables territoriales, que son las de mayor interés en este estudio, se observa que todas ellas afectan a la sobreeducación con los signos esperados. De un lado, residir en un municipio pequeño y, por tanto, acceder a menores oportunidades de empleo eleva la probabilidad de estar sobreeducado. Ofrece especial interés el hecho de que la posibilidad de abarcar un mayor ámbito territorial de búsqueda reduzca de manera significativa dicha probabilidad de sobreeducación. En efecto, la disponibilidad de un vehículo privado y la dotación infraestructural de la provincia (aproximada mediante el número de kilómetros de carretera por vehículo) resultan significativas y con el signo negativo esperado. Así pues, la

---

<sup>13</sup> El número de vehículos y kilómetros de carretera de cada provincia se ha tomado del Anuario Estadístico del INE.

<sup>14</sup> Este resultado queda plenamente confirmado en el meta-análisis realizado por Groot y van den Brink (2000a).

probabilidad individual de sobreeducación vendría parcialmente explicada por factores espaciales, como el tamaño del mercado de trabajo local o la posibilidad de “ampliar” la dimensión del mismo mediante un proceso de búsqueda en mercados de trabajo más alejados gracias a disponer de un vehículo privado que facilite el *commuting* y a una buena red de carreteras.

Estos resultados están en la línea de los obtenidos para el mercado de trabajo alemán por Büchen y Van Ham (2002) y refuerzan la idea de que resulta más importante la dimensión espacial del problema que la denominada sobreeducación diferencial de las mujeres. Es más, los resultados obtenidos apuntarían que las posibles diferencias de sobreeducación entre hombres y mujeres podrían estar determinadas por las dificultades de estas últimas para la movilidad a causa de su mayor responsabilidad en las tareas familiares, especialmente cuando hay hijos de corta edad.

#### **4. CONCLUSIONES**

El objetivo de este trabajo ha consistido en contrastar para la economía española la influencia en la sobreeducación de los mercados de trabajo locales y de otras variables relacionadas con el territorio y con la movilidad espacial.

Para ello, se han utilizado datos individuales de la Encuesta de Presupuesto Familiares 1990-91 y se han aplicado las técnicas econométricas más adecuadas (modelos logit *multilevel* con corrección del sesgo de selección).

Los resultados obtenidos han confirmado que el cuanto mayor es el nivel de estudios mayor es la probabilidad de estar sobreeducado, mientras que esta probabilidad no se ve afectada por la experiencia potencial del individuo en el mercado de trabajo, lo que tendería a rechazar la existencia de sustituibilidad entre diferentes componentes del capital humano.

Un segundo resultado es que hay indicios que en el mercado de trabajo español se podría haber seguido la estrategia de no aceptar un empleo para el que se está sobreeducado a fin de prorrogar la búsqueda de un puesto de trabajo óptimo.

En el caso concreto de las variables relacionadas con el territorio, el contraste de la hipótesis de la sobreeducación diferencial de Frank (1978) que se ha realizado lleva a rechazarla, ya que el riesgo de sobreeducación no es mayor para las mujeres casadas que residen en municipios pequeños. Estos resultados coinciden con la evidencia disponible para otros países.

Con todo, el principal resultado del estudio es que la sobreeducación puede explicarse en parte por factores espaciales, como el menor tamaño del municipio de residencia o la posibilidad de acceder a un mercado de trabajo más amplio al disponer de un vehículo privado que facilite el *commuting* y a la existencia de una suficiente infraestructura viaria. Así pues,

podría afirmarse que quienes padecen una sobreeducación diferencial, más que las mujeres que viven en pareja, son las personas que residen en municipios pequeños y que se encuentran limitadas geográficamente en su ámbito de búsqueda de empleo como resultado de su baja movilidad.

Estos resultados están en línea con los obtenidos para el mercado de trabajo alemán por Büchen y Van Ham (2002) y destacan que resulta más importante la dimensión espacial del problema de la sobreeducación que la denominada sobreeducación diferencial de las mujeres. De hecho, las posibles diferencias de sobreeducación entre hombres y mujeres tal vez se puedan explicar por la menor movilidad espacial de éstas.

Para terminar se apuntan dos líneas de mejora de este estudio sobre las que se está empezando a trabajar. De un lado, podría ser que la disponibilidad de vehículo propio estuviera afectada por la sobreeducación, en la medida que ésta incide negativamente sobre los ingresos salariales, dado que la rentabilidad de los estudios excedentes, aunque positiva, es inferior a la de los estudios requeridos en el puesto de trabajo<sup>15</sup>. Ello podría generar una cierta endogeneidad de la variable “disponer de transporte privado” utilizada en la explicación de la sobreeducación. Un segundo aspecto susceptible de ser mejorado es la medición de la sobreeducación. Dada las limitaciones de la información disponible, debería explorarse la posibilidad de construir medidas basadas en la media más/menos un intervalo distinto a la desviación estándar, o bien basadas en una moda corregida. Disponer de estos diferentes valores de la sobreeducación permitiría, además, realizar un análisis de sensibilidad de los resultados aquí presentados.

## 5. REFERENCIAS BIBLIOGRÁFICAS

- ALBA-RAMÍREZ, A. (1993): “Mismatch in the Spanish Labor Market: Overeducation?”, *The Journal of Human Resources*, 28, pp. 259-278.
- ARTÍS, M., ROMANÍ, J. Y SURIÑACH, J. (2000): “Determinants of Individual Commuting in Catalonia, 1986-91: Theory and Empirical Evidence”, *Urban Studies*, 37 (8): pp. 1431-1450.
- BARRIOS, C. Y MARTÍNEZ, M.A. (1998): *The demand for urban transport: An application of discrete choice model for Cadiz*, XXXVIII Congress of the European Regional Science Association, Vienna, august.
- BLANCO, J. M. (1997): “Comentarios acerca del desajuste educativo en España”, *Papeles de Economía Española*, 72, pp. 117-128.

---

<sup>15</sup> Un panorama de la evidencia internacional al respecto puede verse en Hartog (2000) y en Groot y van den Brink (2000a).

- 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 del IVIE*, WP-EC 96-11.
- BÜCHEL, F. Y BATTU, H. (2002): The Theory of Differential Overqualification: Does it Work?, *IZA Discussion paper*, 511.
- BÜCHEL, F. Y VAN HAM, M. (2002): Overeducation, Regional Labour Markets and Spatial Flexibility, *IZA Discussion paper*, 424.
- CASADO, J.M. (2000): Diferencias de género en los desplazamientos cotidianos por razones laborales, *Documento de Trabajo del IVIE* WP-EC-2000-03.
- DUNCAN, G. Y HOFFMAN, S. (1981): "The economic value of surplus education", *Economics of Education Review*, 1, 1, pp. 75-86.
- FRANK, R. (1978): "Why Women Earn Less: The Theory and Estimation of Differential Overqualification", *American Economic Review*, vol. 68, 3, pp. 360-373.
- FREEMAN, R. B. (1976): *The overeducated American*, Academic Press, Londres.
- GARCÍA MONTALVO, J. (1995): Empleo y sobrecualificación: El caso español, *Documento de Trabajo de FEDEA*, 95-20.
- GARCÍA MONTALVO, J. (2001): "Educación, Empleo y Sobrecualificación", *Economistas*, num.87, año XIX, pp. 282-291.
- GARCÍA SERRANO, C. Y MALO, M. A. (1996): "Desajuste educativo y movilidad laboral en España", *Revista de Economía Aplicada*, vol 4, 11, pp. 105-131.
- GARCÍA SERRANO, C. Y MALO, M. A. (1997): "¿Es diferente el desajuste educativo de las mujeres?", *Información Comercial Española. Revista de Economía*, 760, febrero, pp. 117-128.
- GOLDSTEIN, H. (1995): *Multilevel Statistical Models*, Edward Arnold, Londres.
- GOLDSTEIN, H. Y RASBASH, J. (1996): "Improved approximations for multilevel models with binary responses", *Journal of the Royal Statistical Society, A*, 159, pp. 505-514.
- GROOT, W. Y MAASSEN VAN DEN BRINK, H. (2000a): "Overeducation in the labor market: a meta-analysis", *Economics of Education Review*, 19, pp. 149-158.
- GROOT, W. Y MAASSEN VAN DEN BRINK, H. (2000b): "Skill mismatches in the Dutch labor market", *International Journal of Manpower*, 21, 8, pp. 584-595.
- HARTOG, J. (2000): "Over-education and earnings: where are we, where should we go", *Economics of Education Review*, 19, pp. 131-147.

- HARTOG, J. Y OOSTERBEEK, H. (1988): “Education, allocation and earnings in the Netherlands: Overschooling?”, *Economics of Education Review*, 7, 2, pp. 185-194.
- HECKMAN, J. (1979): “Sample Selection Bias as a Specification Error”, *Econometrica*, 47, pp. 153-161.
- HOX, J. (1998): “Multilevel modeling: When and why?”, en Balderjahn, I., Mathar, R. y Schader, M. (eds.): *Classification, data analysis and data highways*, Springer-Verlag, Nueva York, pp. 147-154.
- KIKER, B. Y SANTOS, M. (1991): “Human capital and earnings in Portugal”, *Economics of Education Review*, 10, 3, pp. 187-203.
- KIKER, B., SANTOS, M. Y MENDES DE OLIVEIRA, M. (1997): “Overeducation and undereducation: evidence for Portugal”, *Economics of Education Review*, 16, 2, pp. 111-125.
- MADRIGAL, M. (2002): *Consideraciones sobre la medición del desajuste educativo*, Comunicación presentada en las XI Jornadas de la Asociación de Economía de la Educación, Lisboa, 27-28 de septiembre de 2002.
- MATAS, A. (1991): “La demanda de transporte urbano: un análisis de las elasticidades y valoraciones del tiempo”, *Investigaciones Económicas*, vol. XV, n. 2, pp. 249-267.
- MCGOLDRICK, K. Y ROBST, J. (1996): “Gender differences in overeducation: a test of the theory of differential overeducation”, *American Economic Review*, 86, pp. 280-285.
- MENDES DE OLIVERA, M., SANTOS, M. Y KIKER, B. (2000): “The role of human capital and technological change in overeducation”, *Economics of Education Review*, 19, pp. 199-206.
- OLIVER, J. Y RAYMOND, J.L. (dirs.) (2002): Educación formal y demanda de cualificación de la mano de obra en España. Una visión a largo plazo, *Centre d’Economia Industrial, Document d’Economia Industrial* n. 14, 136 p.
- ROMANÍ, J., ARTIS, M. Y SURIÑACH, J. (2001): Are Commuting and Residential Mobility Decisions Simultaneous?: the Case of Catalonia (Spain): *Document de Treball de la Divisió de Ciències Jurídiques, Econòmiques i Socials de la Universitat de Barcelona, Colecció d’Economia*, n. E01/73. De próxima aparición en *Regional Studies*.
- RUMBERGER, R. (1987): “The impact of surplus schooling on productivity and earnings”, *Journal of Human Resources*, XXII, (1): pp. 24-50.
- SICHERMAN, N. (1991): “Overeducation in the labour market”, *Journal of Labour Economics*, 9, pp. 101-122.

- SIMPSON, W. (1992): *Urban structure and the labour market: worker mobility, commuting and underemployment in cities*, Clarendon Press, Oxford.
- SINGER, J. D. (1998): “Using SAS PROC MIXED to fit multilevel models, hierarchical models, and individual growth models”, *Journal of Educational and Behavioral Statistics*, 24, 4, pp. 323-355.
- VERDUGO, R. Y VERDUGO, N. (1989): “The impact of surplus schooling on earnings”, *Journal of Human Resources*, 24 (4): pp. 629-643.

**ANEXO 1.**

Tabla 1. Porcentaje de trabajadores sobreeducados según diversos estudios

Autores	Encuesta	Año	Método	Sobreeducados (%)
Alba (1993)	Encuesta de Condiciones de Vida y de Trabajo	1985	Subjetivo (indirecto)	17,00%
García Montalvo (1995)	Encuesta de Población Activa	1985	Objetivo	3,70%
García Montalvo (1995)	Encuesta de Población Activa	1989	Objetivo	6,30%
Oliver y Raymond (dirs.) (2002)	Encuesta de Población Activa	1990	Estadístico (moda)	17,20%
García Serrano y Malo (1996)	Encuesta de Conciencia y Biografía de Clase	1991	Subjetivo (indirecto)	28,40%
García Serrano y Malo (1996)	Encuesta de Conciencia y Biografía de Clase	1991	Subjetivo (directo)	29,40%
García Montalvo (1995)	Encuesta de Población Activa	1991	Estadístico (media)	8,90%
Beneito <i>et al.</i> (1996)	Encuesta de Conciencia y Biografía de Clase	1991	Estadístico (media)	15,90%
García Montalvo (1995)	Encuesta de Población Activa	1993	Objetivo	7,70%
Madrigal (2002)	Panel de Hogares de la Unión Europea	1995	Estadístico (media)	16,50%
Madrigal (2002)	Panel de Hogares de la Unión Europea	1995	Estadístico (moda)	22,80%
Madrigal (2002)	Panel de Hogares de la Unión Europea	1995	Subjetivo (indirecto)	34,70%
	Encuesta de Presupuestos Familiares	1990-91	Estadístico (media)	14,60%

Fuente: Blanco (1997) y elaboración propia.

Tabla 2. Determinantes de la ocupación y de la sobreeducación (primer y segundo paso del procedimiento de Heckman)

Variables	Variable dependiente: ocupado		Variable dependiente: sobreeducado		Variable dependiente: sobreeducado	
	Coefficientes	Error estándar	Coefficientes	Error estándar	Coefficientes	Error estándar
Constante	-0.072	0.263	-25.573	1.212 ***	-25.817	1.243 ***
Mujer (Hombre)	-0.293	0.082 ***				
Mujer* resto ingresos familiares	0.021	0.023				
Mujer con niños (Mujer sin niños)	-0.017	0.119				
Con pareja (Sin pareja)	0.344	0.108 ***				
Mujer con pareja (Mujer sin pareja)	-0.315	0.132 ***				
Mujer casada municipio grande			-0.148	0.270	-0.134	0.272
Mujer soltera municipio grande (Mujer casada municipio pequeño)			-0.495	0.288 *	-0.499	0.289 *
Mujer soltera municipio pequeño			0.012	0.169	-0.014	0.172
Hombre casado municipio grande			-0.529	0.208 ***	-0.460	0.225 **
Hombre soltero municipio grande			-0.152	0.118	-0.077	0.146
Hombre casado municipio pequeño			-0.306	0.272	-0.241	0.285
Hombre soltero municipio pequeño			-0.046	0.169	0.000	0.177
Hijo que convive con padres (no hijo que convive con padres)	-0.350	0.087 **	0.161	0.151	0.057	0.192
Experiencia	0.061	0.010 ***	-0.016	0.016	-0.003	0.022
Experiencia al cuadrado	-0.001	0.000 ***	0.000	0.000	0.000	0.000
Estudios	0.053	0.014 ***	1.448	0.040 *	1.458	0.042 ***
Trabajo a tiempo parcial (Trabajo a tiempo completo)			-0.287	0.324	-0.299	0.326
Tasa de paro provincial	-0.021	0.006 ***	-0.005	0.010	-0.009	0.011
Lambda					0.872	0.990
*** 1%, ** 5%, *10%						

Tabla 3. Determinantes de la ocupación y de la sobreeducación (primer y segundo paso del procedimiento de Heckman)

Variables	Variable dependiente: ocupado		Variable dependiente: sobreeducado		Variable dependiente: sobreeducado	
	Coefficientes	Error estándar	Coefficientes	Error estándar	Coefficientes	Error estándar
Constante	0.277	0.504 ***	-24.294	0.788 *	-25.004	0.887
Mujer (Hombre)	-0.438	0.143	0.074	0.117	-0.085	0.147
Mujer* resto ingresos familiares	0.011	0.040				
Mujer con niños (Mujer sin niños)	0.052	0.223	-0.157	0.185	-0.166	0.186
Con pareja (Sin pareja)	1.271	0.186 ***				
Mujer con pareja (Mujer sin pareja)	-0.845	0.260 ***	0.176	0.180	0.281	0.190
Experiencia	0.123	0.018 ***	-0.026	0.015 *	0.003	0.022
Experiencia al cuadrado	-0.002	0.001 ***	0.000	0.000	0.000	0.000
Estudios	0.102	0.025 ***	1.453	0.041 ***	1.473	0.042 ***
Trabajo a tiempo parcial (Trabajo a tiempo completo)			-0.353	0.324	-0.368	0.327
Disponer de transporte privado (no disponer de transporte privado)			-0.728	0.155 ***	-0.722	0.155 ***
Residir en un municipio pequeño (residir en un municipio grande)			0.409	0.156 ***	0.397	0.159 ***
Tasa de paro provincial	-0.041	0.013 ***	-0.013	0.010	-0.018	0.011
Km carreteras/automóviles	-0.031	0.380	-0.558	0.303 **	-0.533	0.311 *
Lambda					1.222	0.680 *
*** 1%, ** 5%, *10%						