

XIII Encuentro de Economía Aplicada

Sevilla, 10-11 de junio de 2010

Diferencias de género en el desajuste educativo. El papel del capital humano y de las características empresariales. *

(versión –muy- preliminar)

Raúl Ramos

AQR-IREA, Universitat de Barcelona, rmos@ub.edu, Telf. 934024310 - Fax 934021821

Departamento de Econometría, Estadística y Economía Española, Avda. Diagonal 690, 08034 Barcelona

Elisabet Motellón

Universitat Oberta de Catalunya (UOC) y AQR-IREA, Universitat de Barcelona, emotellon@uoc.edu,

Telf. 932537501 - Fax 934176495 Avda. Tibidabo 39-43, 08035 Barcelona

Enrique López-Bazo

AQR-IREA, Universitat de Barcelona, elopez@ub.edu, Telf. 934037041 - Fax 934021821

Departamento de Econometría, Estadística y Economía Española, Avda. Diagonal 690, 08034 Barcelona

Resumen

Estudios recientes han señalado que las diferencias de género en el mercado de trabajo español están relacionadas con la elevada segregación ocupacional de las mujeres en trabajos de baja calificación que contrasta con el superior nivel educativo de las mujeres en relación a los hombres. Desde esta perspectiva, uno de los factores que puede incidir con mayor importancia en la existencia de diferencias salariales entre hombre y mujer es el desajuste educativo. El objetivo del presente trabajo consiste en analizar la diferencia de género en la propensión a ocupar un puesto de trabajo que requiere un nivel educativo distinto al acreditado por el trabajador, prestando especial atención a, por una parte, el efecto atribuible a las características del trabajador y, por otra, a las correspondientes al puesto de trabajo y al establecimiento. Para ello, en primer lugar se cuantifica la incidencia del desajuste educativo entre ambos colectivos aplicando el método estadístico a partir de los microdatos de la Encuesta de Estructura Salarial (EES) para 1995 y 2006 y, en segundo lugar, se aplican técnicas de descomposición detallada de la diferencia de género en la probabilidad de estar sobre e infraeducado. Los resultados obtenidos confirman la magnitud de la brecha de género en tasas de desajuste educativo, tanto por sobre como por infraeducación, que no se mantienen estables a lo largo del periodo analizado y en las que la contribución de las diferencias en características asociadas al puesto de trabajo y al establecimiento ha jugado un papel destacado.

Palabras clave

Diferencias de género, sobre e infraeducación, modelos probabilísticos, descomposición detallada

Códigos JEL

J31, J24, J71, C25

* Los autores desean agradecer el apoyo recibido del Ministerio de Educación y Ciencia mediante el proyecto ECO2008-05314/ECON.

1. Introducción

Según datos de la Encuesta Europea de Estructura Salarial para 2006 publicados por Eurostat¹, el diferencial salarial por sexo en España se situaba en el 17,9%, un valor inferior al observado en 2002 (20,2%) pero superior al de la mayoría de países de nuestro entorno (Irlanda, 17,2%; Suecia, 16,5%; Luxemburgo, 10,7% Francia, 15,8%; Portugal, 8,4% o Italia, 4,4%). Los análisis existentes sobre el origen de dicho diferencial han mostrado que éste no obedece mayoritariamente a las diferencias en las características observables entre hombres y mujeres (De la Rica y Ugidos, 1995), y que es mayor para las mujeres más cualificadas (De la Rica et al., 2008). También han concluido que se trata de un fenómeno asociado a la segregación laboral de las mujeres (Simón et al., 2008). En la práctica, éstas tienden a estar empleadas en mayor medida en ocupaciones y establecimientos que pagan salarios comparativamente bajos a trabajadores con las mismas características productivas, de modo que una porción significativa del diferencial salarial por razón de sexo se debe a su segregación en segmentos del mercado de trabajo mal remunerados (Amuedo-Dorantes y De la Rica, 2006) y de baja calificación que contrasta, además, con el superior nivel educativo de las mujeres en relación a los hombres.

Así pues, desde esta perspectiva, uno de los factores que puede incidir con mayor importancia en la existencia de diferencias salariales entre hombre y mujer es el desajuste educativo. El desajuste educativo es aquella situación en que el nivel educativo del trabajador difiere del requerido para desempeñar las tareas asociadas a su puesto de trabajo. En concreto, puede suceder que un trabajador tenga un nivel educativo superior al necesario, en cuyo caso se diría que dicho trabajador está sobreeducado, mientras que en caso contrario estaría infraeducado. La literatura sobre el tema ha confirmado el impacto negativo del desajuste educativo sobre los salarios: aunque los años de estudios sobrantes tienen un impacto positivo sobre el salario, su rendimiento es claramente inferior al de los años requeridos, mientras que un trabajador infraeducado cobra menos que sus compañeros de ocupación que poseen los estudios requeridos (Groot y Maasen van den Brink, 2000 y Ng, 2003)

Teniendo en cuenta la importancia del desajuste educativo como determinante de las diferencias salariales entre hombres y mujeres, el objetivo del trabajo consiste en analizar la diferencia de género en la propensión a ocupar un puesto de trabajo que requiere un nivel educativo distinto al acreditado por el trabajador. Para ello, en primer lugar se cuantifica la incidencia del desajuste

¹ <http://epp.eurostat.ec.europa.eu/tgm/table.do?tab=table&init=1&language=en&pcode=tsiem040&plugin=1>

educativo entre ambos colectivos aplicando el método estadístico (el único posible con los datos que se utilizarán); en segundo lugar, se especifican y estiman modelos de elección discreta orientados a identificar los determinantes del desajuste educativo y, por último, se aplican técnicas de descomposición detallada de la diferencia de género en la probabilidad de estar sobre e infraeducado.

El estudio utiliza los microdatos de la Encuesta de Estructura Salarial (EES) para 1995 y 2006. Esta base de datos, además de incorporar información muy rica sobre salarios y características individuales, permite analizar la influencia de distintas características de los establecimientos sobre la diferencia de género en la propensión de que un trabajador presente desajuste educativo, un aspecto que no ha sido estudiado hasta el momento en el caso español y únicamente de manera marginal a nivel internacional. Para ello, y desde un punto de vista metodológico, se aplica la descomposición detallada en modelos no lineales propuesta en Yun (2004) a las diferencias en la probabilidad media de desajuste educativo entre hombres y mujeres. De esta forma, además de determinar la contribución a las diferencias de la dotación de características y de los parámetros asociados a las mismas (su impacto) en términos globales, es posible aislar la contribución asociada a las características personales del trabajador (básicamente su capital humano) y la vinculada con las características de su puesto de trabajo y del establecimiento en el que presta sus servicios.

El resto del trabajo se estructura en cuatro apartados. En primer lugar, se realiza una breve revisión de la literatura sobre desajuste educativo y diferencias de género; a continuación, y en segundo lugar, se presenta la base de datos utilizada en el ejercicio empírico y se define la variable objeto de análisis. En el cuarto apartado se describe la evidencia empírica obtenida, tanto por lo que respecta a la estimación de los modelos probabilísticos para ambos géneros, como a los resultados correspondientes a la descomposición de las diferencias en la probabilidad de desajuste educativo. Por último, el trabajo finaliza resumiendo las principales conclusiones y apuntando algunas líneas futuras de investigación.

2. Revisión de la literatura

El fenómeno del desajuste educativo fue analizado por primera vez por Freeman (1976). Sin embargo, dicho concepto no pasó a formar parte de la corriente principal de la economía laboral hasta la publicación del trabajo de Duncan y Hoffman (1981). A partir de las bases fijadas por estos autores, la relación entre nivel de estudios y empleo ha sido analizada desde esta perspectiva en

diversos países desarrollados (Groot y Maassen van den Brink, 2000). El caso español no ha sido distinto y a partir del trabajo de Alba-Ramírez (1993), diversos autores (García Montalvo, 1995 o García-Serrano y Malo, 1996, entre otros) han analizado la importancia de este fenómeno para la economía española.

Sin embargo, un aspecto que ha merecido escasa atención en la abundante literatura sobre el tema es la existencia de diferencias en la incidencia del desajuste educativo entre hombres y mujeres. El trabajo pionero en este ámbito es el de Frank (1978) para Estados Unidos. Dicho estudio tenía como objetivo analizar la influencia de la sobreeducación sobre el diferencial salarial entre hombres y mujeres. Para ello, Frank (1978) desarrolla un modelo donde supone que los individuos buscan trabajo tratando de maximizar sus ingresos, lo que implica minimizar la diferencia entre sus estudios y los requerimientos educativos de la vacante ya que así conseguiría maximizar el rendimiento de su inversión en capital humano. 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 dentro del mercado de trabajo local en que se ha establecido el matrimonio, que es el que minimiza la sobreeducación del marido. Así pues, las posibilidades de la esposa de obtener un *match* óptimo serían claramente inferiores a las del marido y, por tanto, el mayor desajuste educativo de las mujeres casadas explicaría (al menos, parcialmente) el diferencial salarial observado entre hombres y mujeres. Los resultados obtenidos por Frank (1978) a partir de datos para la economía americana para el año 1967 confirman dicha hipótesis.

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 con datos estadounidenses para 1985 y, en cambio, no encuentran evidencia de sobreeducación diferencial. Büchel y Battu (2003) llegan a una conclusión similar a partir del análisis de datos de 1995 para la economía alemana.

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 estas diferencias no parecen estar explicadas por el tamaño del mercado de trabajo local).

Sanromá y Ramos (2004) también analizan la existencia de sobreeducación diferencial para el caso español pero utilizando los microdatos de la Encuesta de Presupuestos Familiares. Sus conclusiones están en línea con los encontrados por García-Serrano y Malo (1997): a pesar de que se observa una mayor incidencia de la sobreeducación entre las mujeres, ésta no parece estar relacionada con características del mercado de trabajo local que sólo les afecten a ellas. De hecho, sus resultados apuntan a qué quienes padecen una sobreeducación diferencial (más que las mujeres que viven en pareja) son las personas (tanto hombres como mujeres) que residen en municipios pequeños y que se encuentran limitadas geográficamente en su ámbito de búsqueda de empleo como resultado de su baja movilidad.

Johanson y Katz (2007) analizan el impacto de distintas características individuales sobre el diferencial entre hombres y mujeres en términos de desajuste educativo. A diferencia de los autores anteriores, no centran su interés en la hipótesis de la sobreeducación diferencial sino que su objetivo es identificar cuáles son las características individuales que influyen sobre la distinta probabilidad de estar sobre o infra-educado para hombres o mujeres. Sus resultados muestran que la educación y la experiencia afectan de manera distinta al desajuste educativo de hombres y mujeres. Tanto hombres como mujeres tienen una menor probabilidad de estar sobreeducados si están casados, pero el efecto es muy superior para los hombres. En cambio, sus resultados no apoyan la hipótesis de que las restricciones derivadas de estar casadas o tener hijos pequeños constituyen la principal explicación de la mayor sobreeducación de las mujeres.

Robst (2007) realiza una aproximación diferente al estudio de las diferencias de género en lo que se refiere al desajuste educativo. En concreto, a partir de una encuesta específica para Estados Unidos realizada en 1993, analiza las respuestas de una muestra de individuos que experimentan desajuste educativo sobre cuáles son las razones por las que se encuentran en dicha situación. Sus resultados muestran que los motivos que argumentan hombres y mujeres son muy distintos. En concreto, para los hombres, las razones para aceptar un trabajo que requiere una formación distinta a la que tienen están relacionadas con las mayores posibilidades de promoción, mientras que, en cambio, las mujeres muestran su preferencia por unas condiciones laborales más flexibles, la localización de su trabajo y motivos familiares.

El análisis de Robst (2007) muestra, por tanto, que al margen de la situación familiar y de las restricciones geográficas a las que se enfrentan las mujeres, puede haber otras características relacionadas con el puesto de trabajo o con las condiciones laborales que contribuyan a explicar el mayor desajuste educativo de las mujeres. El objetivo de este trabajo es contrastar la importancia relativa de estas características empresariales y del puesto de trabajo y analizar su capacidad explicativa sobre el diferencial de género en la probabilidad de experimentar desajuste educativo. Hasta dónde nosotros conocemos, el único trabajo que ha considerado la influencia relativa de este tipo de características sobre la probabilidad de desajuste educativo (aunque sin analizar las diferencias hombre-mujer) es el de Karakaya et al. (2007). Dichos autores utilizan los datos de la Encuesta de Estructura Salarial de 1995 para Bélgica para cuantificar la influencia de distintos factores relacionados con la empresa sobre la probabilidad de sobreeducación. Sus resultados muestran la importancia de algunas de estas características como, por ejemplo, la titularidad pública o privada de la empresa o el tipo de contrato de los trabajadores.

Así pues, la evidencia internacional ha puesto de manifiesto la importancia de las características personales, pero también de las características del puesto de trabajo y de la empresa donde trabajan los individuos. Además, la evidencia para el caso español no es concluyente en lo que se refiere a si existen o no diferencias significativas entre hombres y mujeres. Mientras que Marzo-Navarro (2007) encuentran una mayor sobreeducación entre las mujeres en relación a los hombres, otros estudios como García-Serrano y Malo (1996) o Aguilar y Navarro (2003) no detectan una diferencia significativa entre ambos sexos. En el próximo apartado se analizan los datos de la Encuesta de Estructura Salarial para 1995 y 2006 con el objetivo de aportar nueva evidencia empírica en este ámbito.

3. Base de datos y variables

3.1. La Encuesta de Estructura Salarial

Los microdatos utilizados en este trabajo provienen de la Encuesta de Estructura Salarial (desde aquí EES) para 1995 y 2006. El diseño de la EES corresponde a un amplio muestreo en dos etapas de trabajadores a partir de las cuentas de cotización de las empresas en la Seguridad Social. La EES abarca únicamente a empleados por cuenta ajena y cubre un amplio abanico de sectores productivos. La principal característica distintiva de la EES en relación a otras operaciones estadísticas similares es que facilita datos emparejados de empresas y trabajadores, una información que permite considerar apropiadamente la influencia de las características empresariales sobre el

desajuste educativo de los trabajadores, que constituye el objetivo principal del trabajo.² La EES contiene una variada información sobre los trabajadores (salario, sexo, edad, educación, antigüedad, ocupación, tipo de contrato y tipo de jornada) y sobre sus empresas (sector, tamaño, tipo de convenio colectivo y comunidad autónoma). Con el fin de tener una cobertura empresarial homogénea, en la EES 2006 se han filtrado las observaciones de los sectores de actividad no cubiertos por la EES 1995 y se han eliminado las observaciones relativas a los individuos que trabajan en establecimientos de menos de diez trabajadores. Las muestras finales resultantes tienen un tamaño de 132.684 observaciones para 1995 y de 136.261 observaciones para 2006. En el anexo 1 se presentan los principales descriptivos de dichas muestras tanto para el conjunto de los individuos analizados como para hombres y mujeres por separado.

3.2. La medición del desajuste educativo

La literatura ha desarrollado tres métodos distintos para medir el desajuste educativo (Hartog, 2000): el método objetivo, el método subjetivo (directo e indirecto) y el método estadístico.

El método objetivo se basa en el análisis realizado por analistas laborales de los puestos de trabajo y su traslación a requisitos educativos. Una vez establecidos estos requisitos se comparan con la formación que poseen los individuos que los ocupan a fin de determinar el grado de ajuste. Las limitaciones de este método residen en la simplificación derivada de reducir a unos pocos niveles educativos las múltiples características de los empleos existentes y, especialmente, en la dificultad de elaborar los mencionados requisitos, lo que provoca su falta de actualización. Todo ello explica el reducido número de trabajos que aplican este método: Rumberger (1987) en Estados Unidos a partir del *Dictionary of Occupational Titles*; Hartog y Oosterbeek (1988) en Holanda; Kiker y Santos (1991) y Kiker *et al.* (1997) en Portugal y -en una adaptación específica para el mercado de trabajo español- García Montalvo (1995) para España.

El método subjetivo se basa en las respuestas que ofrecen los individuos a preguntas relacionadas con su nivel de formación y los requerimientos del puesto de trabajo que ocupan. Se suele distinguir una vía directa, consistente en solicitar al individuo que se autoclasifique como adecuadamente educado, infraeducado o sobreeducado, y una vía indirecta, en que se clasifica al individuo tras comparar su nivel de estudios con la respuesta relativa a la formación necesaria para desempeñar su puesto de trabajo. Los resultados obtenidos por ambos procedimientos -directo e indirecto- no

² No obstante, cabe indicar como contrapartida que la ausencia de información sobre no ocupados en dicha encuesta impide controlar la posible existencia de selección muestral en el análisis empírico realizado en este trabajo.

acostumbran a coincidir, dado que algunos trabajadores son reacios a clasificarse como infraeducados.

El método subjetivo no requiere de tanta información como el anterior, pero se basa únicamente en la percepción del propio individuo respecto a su trabajo, lo que limita su rigor. La facilidad de aplicación de este método ha impulsado una extensa utilización del mismo en múltiples estudios. Algunos trabajos que emplean este método son Duncan y Hoffman (1981) y Sicherman (1991) para Estados Unidos, Hartog y Oosterbeek (1988) para Holanda y Alba-Ramírez (1993) y García-Serrano y Malo (1996) para España.

El método estadístico consiste en comparar los años de estudio de un individuo con los años de estudio de los trabajadores empleados en su misma ocupación. Se considera sobreeducado (infraeducado) a aquel trabajador cuyos años de educación superan en más (menos) de una desviación estándar los años medios de educación en la ocupación que realiza (Verdugo y Verdugo, 1989). Kiker *et al.* (1997) han propuesto el uso de la moda en lugar de la media, argumentando que el nivel de educación adecuado para desempeñar una ocupación es aquel que predomina entre los trabajadores de la ocupación. Sin embargo, tal y como señalan Mendes de Olivera *et al.* (2000), para poder aplicar este método es recomendable que la moda abarque al menos un 60% de los trabajadores de la ocupación analizada. Algunos estudios para el caso español que han utilizado este método para el caso español son Beneito *et al.* (1996), Oliver y Raymond (2002) y Sanromá y Ramos (2004), entre otros.

Tanto en su versión media como moda, este método tiene la ventaja de necesitar únicamente para su cálculo información sobre nivel educativo y ocupación -suficientemente desagregada- del individuo, lo cual evita los problemas de falta de actualización y subjetividad, respectivamente, de los métodos objetivo y directo. Con todo, este método estadístico tiene como debilidad su carácter parcialmente endógeno, ya que el nivel educativo de los actuales trabajadores no sólo depende de los requerimientos del puesto de trabajo, sino también de las políticas de empleo de las empresas y de las características de la oferta de cualificaciones (Hartog, 2000).

Lamentablemente, los diferentes métodos ofrecen resultados muy diversos aun cuando se apliquen a una misma base de datos (Groot y Maassen van den Brink, 2000). 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, 2002). 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 EES no contiene información subjetiva, pero informa con detalle sobre el nivel de estudios y la ocupación desempeñada, en este trabajo se ha utilizado el método estadístico para obtener una medida del desajuste educativo para 1995 y para 2006. La EES presenta un elevado nivel de detalle en lo que se refiere a los niveles educativos de los trabajadores pero un desglose de ocupaciones de 2 dígitos de la Clasificación Nacional de Ocupaciones inferior al utilizado habitualmente en la literatura (3 dígitos). La diferencia es importante pues una desagregación de 3 dígitos implica disponer de información sobre más de 200 ocupaciones distintas mientras que la desagregación de 2 dígitos reduce dicho número a algo más de 60. Para solucionar este problema se ha optado por introducir en el cálculo de los requerimientos de formación la dimensión sectorial. En concreto, y teniendo en cuenta que los requerimientos educativos para algunas ocupaciones pueden ser muy distintos entre empresas de distintos sectores se ha optado por analizar algunas ocupaciones de manera independiente para la industria, la construcción y los servicios. A continuación, y con el objetivo de aplicar el criterio de la moda propuesto por Mendes de Olivera *et al.* (2000), es decir que la moda englobe el 60% del total de individuos analizados en un número de elevado de categorías, se han agrupado algunas ocupaciones con aquellas más cercanas. El número total de ocupaciones consideradas ha sido superior a 150.

Los resultados obtenidos de manera independiente para cada una de las muestras analizadas (1995 y 2006) se muestran en la tabla 1. Tal y como se puede observar a partir de dicha tabla, el porcentaje de trabajadores adecuadamente educados se ha reducido entre los dos años considerados, pasando del 48% en 1995 al 38% en 2006. Esta reducción ha ido acompañada tanto de un incremento de los trabajadores infraeducados (que han pasado de representar el 20% al 32%) como de una cierta estabilidad en la importancia relativa de los trabajadores sobreeducados que han pasado del 32% en 1995 al 30% en 2006. Si nos centramos en los resultados desagregados por sexos, existen claras diferencias entre 1995 y 2006. Mientras que en 1995, la infraeducación afectaba en mayor medida a las mujeres y la sobreeducación a los hombres, la situación se ha invertido en 2006: un 32% de las mujeres y un 29% de los hombres están sobreeducados y un 29% de las mujeres y 34% de los hombres están infraeducados. Esta situación es mucho más cercana a la observada en la mayoría de economías occidentales (Groot y Maasen van den Brink, 2000). En consecuencia, la diferencia en la tasa de sobreeducación de hombres y mujeres se situó en 5 puntos porcentuales en 1995 y en -2,8 puntos en 2006, mientras que la correspondiente a la tasa de infraeducación estuvo en -3,6 puntos y en 4,6 puntos en esos mismos años. Sin embargo, es posible que las diferencias de género observadas se deban simplemente a diferencias en características observables entre hombres y mujeres (véase la tabla del Anexo). Es decir, que esas diferencias en lo que podemos calificar como probabilidad bruta de desajuste educativo se desvanezcan al controlar por las diferencias de género

en las características que determinan ese desajuste, de forma que no se observen diferencias en las probabilidades netas.

En los siguientes apartados se especifican y estiman modelos probabilísticos que incluyen las características personales y las empresariales y del puesto de trabajo, para valorar la posibilidad de que, además, existan diferencias de género en el impacto de esas características. Y se realiza la descomposición de las diferencias en las probabilidades de hombres y mujeres para determinar la contribución de las diferencias en dotación de características y en el impacto de las mismas. Una contribución relevante de estas últimas deberá ser interpretada como evidencia a favor de la presencia de diferencias de género en la probabilidad neta de desajuste educativo.

4. Resultados

4.1. Los determinantes del desajuste educativo para hombres y mujeres

La magnitud de la brecha entre hombres y mujeres en las tasas brutas de desajuste educativo, así como su evolución, nos conduce a plantearnos su origen. Como se ha comentado, el origen de las diferencias de género en los niveles de sobre e infraeducación puede radicar en las discrepancias en las características personales y/o laborales y, además, en las diferencias en el impacto de las mismas. Para cuantificar la contribución de ambos factores a la explicación de la diferencia en la probabilidad de desajuste educativo por género emplearemos la metodología propuesta en Yun (2004). Para ello, en primer lugar, se ajusta un modelo probit al evento de padecer desajuste educativo por sobreeducación y otro por infraeducación, bajo el supuesto de que la probabilidad de cada uno de esos eventos está relacionada con las características personales de los trabajadores, básicamente su dotación de capital humano, y con el conjunto de características observables de los puestos de trabajo que desempeñan y de las empresas para las cuales prestan sus servicios.³

$$prob(DE)_G = \Phi(X_G \beta_G) \quad (1)$$

donde $prob(DE)$ denota la probabilidad de tener desajuste educativo, sobre o infraeducación según el caso, Φ la función de distribución acumulada de la normal estándar, X representa el conjunto de características observables y β es el vector de coeficientes. El subíndice $G=H, M$ denota el género

³ Como alternativa a la especificación de modelos probabilísticos independientes para cada una de estas dos situaciones se estimó un modelo logit multinomial. Los resultados, disponibles para el lector interesado, coinciden en esencia con los obtenidos para los modelos probit aquí descritos.

de los individuos, hombres o mujeres respectivamente. La estimación de esos coeficientes en cada colectivo será utilizada, en una etapa posterior, para la obtención de la descomposición de las diferencias de género en prob(DE) .

Para aproximar el efecto de las distintas características observadas sobre la probabilidad de desajuste educativo en ambos géneros y, para valorar los posibles cambios acontecidos en los efectos entre 1995 y 2006, se han estimado distintos modelos probit para cada uno de los años considerados y tanto para el conjunto de individuos analizados como para hombres y mujeres de manera separada. Además, en cada uno de esos casos se han estimado tanto modelos que analizan la probabilidad de estar sobreeducado en relación a estar adecuadamente educado o infraeducado como modelos orientados al análisis de la probabilidad de estar infraeducado.⁴

Las variables explicativas de estos modelos se pueden agrupar en 3 bloques: un primer bloque relacionado con el capital humano del individuo (estudios, experiencia y antigüedad en la empresa); un segundo bloque relacionado con las características del puesto de trabajo (tipo de jornada, tipo de contrato y ocupación⁵) y, por último, un tercer bloque que recoge las características de la empresa (sector de actividad, número de trabajadores, participación pública, mercado principal, tipo de convenio, y comunidad autónoma donde se ubica).

Los resultados de estimar estos distintos modelos se presentan en las tablas 2 y 3. En concreto, en la tabla 2 se presentan los efectos marginales evaluados en la media sobre la probabilidad de estar sobreeducado para cada una de las variables analizadas, mientras que en la tabla 3 se presentan los efectos marginales sobre la probabilidad de estar infraeducado. Tal y como se puede observar a partir de estos resultados, una vez controladas las distintas características observables, los hombres presentan una menor probabilidad de sobreeducación (y mayor de infraeducación) que las mujeres. En ambos casos, el diferencial entre hombres y mujeres ha aumentado entre 1995 y 2006. La probabilidad de que un hombre estuviese sobreeducado era 1,45 puntos inferior a la de una mujer en 1995, mientras que en 2006 esa misma probabilidad es de 3,15 puntos inferior. La probabilidad de que un hombre estuviese infraeducado era 1,08 puntos superior a la de una mujer en 1995 y se ha situado en 7,13 puntos superior en 2006. Se esperaría, pues, que esta mayor incidencia del desajuste educativo sobre las mujeres (una vez controlado el efecto de las características observables) tuviese

⁴ Alternativamente se estimaron modelos en los que la alternativa a la sobre o a la infraeducación era la adecuada educación. En esencia los resultados en esos casos fueron similares a los aquí presentados.

⁵ Nótese que la inclusión de la ocupación en la lista de regresores podría causar un problema de endogeneidad al utilizarse dicha variable para la definición de la sobre y de la infraeducación. No obstante, en este ejercicio ese problema queda claramente minimizado al estimar un modelo probabilístico asociado a esos fenómenos (y no al grado de incidencia del mismo) y al haber combinado ocupaciones y sectores en la definición de esas variables de interés.

un efecto negativo sobre la evolución del diferencial salarial entre hombres y mujeres en el mercado de trabajo español.

Un segundo resultado a destacar es que los estudios aumentan la probabilidad de estar sobreeducado (y disminuyen la de estar infraeducado). Éste es un resultado usual en la literatura, tanto internacional como española. Un aspecto a destacar es que los años de estudios penalizan menos a los hombres que a las mujeres. En 2006 la probabilidad de estar sobreeducado aumenta en 12 puntos por año de estudio para las mujeres mientras que para los hombres sólo lo hace en 5 puntos. En 1995 prácticamente no existían diferencias entre sexos.

En lo que se refiere a la experiencia potencial y a la antigüedad, los resultados no confirman la existencia de una cierta sustituibilidad entre educación y otras formas de capital humano. En concreto, cabría esperar que la experiencia potencial afectase negativamente a la probabilidad de estar sobreeducado y positivamente a la infraeducación. Sin embargo, los resultados son diversos tanto para 1995 como para 2006 y también entre hombres y mujeres por lo que no existe una conclusión clara en lo que se refiere al efecto de ambas variables.

Entrando ya a comentar los bloques de variables relacionados con las características de los puestos de trabajo y con las características empresariales, trabajar a jornada completa o disponer de un contrato indefinido reducen la probabilidad tanto de estar sobreeducado como infraeducado. Cabe destacar, sin embargo, que el coeficiente asociado a la variable relativa al tipo de contrato no resulta estadísticamente significativo para las mujeres en 2006. El grupo de variables ficticias relacionado con las ocupaciones muestra que en las profesiones con mayor requerimientos educativos la probabilidad de estar sobreeducado (infraeducado) es menor (mayor) sin que existen diferencias importantes ni entre años ni entre sexos. Un resultado similar se encuentra para el grupo de variables ficticias relativo al sector de actividad de la empresa.

Trabajar en empresas con un elevado número de trabajadores reduce la probabilidad de estar infraeducado y únicamente para los hombres y en 1995 la probabilidad de estar sobreeducado. El primer resultado está en línea con lo esperado (y con la literatura internacional) ya que cabe esperar que en una empresa grande las posibilidades para un trabajador de realizar un mejor *match* en el “mercado interno de la empresa” serán superiores que en una empresa pequeña. En cambio, el segundo resultado resulta sorprendente. Lo mismo ocurre con los trabajadores de las empresas con participación pública quienes experimentan menores niveles de infraeducación (lo esperado), pero, en cambio, mayores niveles de sobreeducación. Los resultados relativos a la variable relacionada

con el principal mercado de la empresa también son similares: en las empresas orientadas principalmente a mercados locales, la incidencia de la infraeducación es inferior pero la de la sobreeducación es superior. Por último, y en lo que se refiere a los distintos tipos de convenio colectivo por los que pueden verse afectados los trabajadores, los resultados tampoco son concluyentes: en algunos casos, una mayor centralización de la negociación parece reducir el desajuste educativo (para los hombres en 2006), pero en otros, los resultados señalan en la dirección contraria (por ejemplo, para las mujeres sobreeducadas en 1995).

En cualquier caso, las estimaciones de los modelos para hombres y mujeres revelan diferencias sustanciales en los efectos marginales tanto de las distintas componentes del capital humano como de las características de puesto de trabajo y empresa. Ello confirma la posibilidad de que las diferencias en las tasas de desajuste educativo puedan ser, al menos en parte, ocasionadas por diferencias en el impacto de esas características y no únicamente de una distinta dotación de las mismas en hombres y mujeres. La descomposición en el siguiente subapartado nos permitirá valorar esta circunstancia.

4.2. Descomposición detallada de las diferencias de genero en la probabilidad de desajuste educativo

El método generalizado de descomposición de diferencias en el primer momento propuesto en Yun (2004) nos proporciona una metodología válida para el caso de formas funcionales no lineales como la del modelo probabilístico en (1). El punto de partida es la siguiente expresión para la diferencia de probabilidades de desajuste educativo entre el grupo hombres (H) y mujeres (M):

$$\overline{prob(DE)}_H - \overline{prob(DE)}_M = [\overline{\Phi(X_H\beta_H)} - \overline{\Phi(X_M\beta_H)}] + [\overline{\Phi(X_M\beta_H)} - \overline{\Phi(X_M\beta_M)}] \quad (2)$$

Esta diferencia puede explicarse a través de la suma de dos términos. El primero recoge el efecto atribuible a las diferencias en las características de los colectivos, mientras que el segundo cuantifica la contribución asignable a las diferencias de rendimientos de las mismas. Esta segunda componente nos estaría indicando que ante igualdad de características, el impacto de éstas sobre la probabilidad de padecer desajuste educativo es distinta en función del sexo del individuo.

Pero a partir de esta especificación en (2) no podemos detallar la contribución de cada variable, o conjunto de variables, a la diferencia de probabilidades. Y ésta es una cuestión importante en este trabajo dado el interés en valorar la contribución de, por una parte, las diferencias de género

asociadas al capital humano y, por otra, a los puestos de trabajo y a los establecimientos. Para resolver esta cuestión, y obtener un mayor detalle de la diferencia total, Yun (2004) sugiere realizar una transformación, en dos etapas, de la especificación general en (2). La primera consiste en evaluar la función en los valores medios de los regresores, mientras que en la segunda se aplica una expansión de Taylor de primer orden con el objetivo de linealizar la forma funcional. Como resultado se obtiene una expresión equivalente a la anterior:

$$\overline{prob(DE)}_H - \overline{prob(DE)}_M = \sum_{i=1}^k P_{\Delta X}^i [\overline{\Phi(X_H \beta_H)} - \overline{\Phi(X_M \beta_H)}] + \sum_{i=1}^k P_{\Delta \beta}^i [\overline{\Phi(X_M \beta_H)} - \overline{\Phi(X_M \beta_M)}] \quad (3)$$

donde:

$$P_{\Delta X}^i = \frac{(\overline{X}_H^i - \overline{X}_M^i) \beta_H^i}{(\overline{X}_H - \overline{X}_M) \beta_H}, \quad P_{\Delta \beta}^i = \frac{\overline{X}_M^i (\beta_H^i - \beta_M^i)}{\overline{X}_M (\beta_H - \beta_M)}$$

$$\sum_{i=1}^k P_{\Delta X}^i = \sum_{i=1}^k P_{\Delta \beta}^i = 1$$

$P_{\Delta X}^i$ y $P_{\Delta \beta}^i$ denotan, respectivamente, los pesos correspondientes a las variables y a su rendimiento. Por tanto, estos pesos permiten distribuir la contribución de las diferencias totales en características y en rendimientos entre las distintas variables observables.

En resumen, esta descomposición de las diferencias en el caso de modelos no lineales nos permite identificar si las discrepancias en la probabilidad de desajuste educativo, tanto para el caso de la sobre como de la infraeducación, entre hombres y mujeres proceden del hecho de tener una dotación diferente de capital humano y estar empleados en puestos de trabajo y empresas distintas o si, por el contrario, son fruto de diferencias en la probabilidad neta de desajuste educativo por razón de género. Sólo de confirmarse este último supuesto podremos afirmar que hombres y mujeres difieren realmente en sobreeducación y/o infraeducación, al hacerse un uso distinto de su capacidad derivada del nivel educativo alcanzado, confirmándose así las diferencias de género. La citada descomposición nos permitirá, además, detallar qué tipo de variables contribuyen en mayor medida a su constitución y, en consecuencia determinar la contribución de las características del puesto de trabajo y el establecimiento, en términos absolutos y en relación a las correspondientes al capital humano.

Los resultados de la descomposición detallada de la diferencia de género en la probabilidad de desajuste educativo para los años 1995 y 2006 se sintetizan en la Tabla 4, para el estudio de la sobreeducación, y en la Tabla 5, para el caso de la probabilidad de estar infraeducado. Cabe indicar

que, dado que entre el conjunto de determinantes de la probabilidad aparecen un número importante de variables ficticias, hemos aplicado la restricción de identificación propuesta por Gardeazabal y Ugidos (2004). Esta restricción permite obtener estimaciones de la contribución de cada categoría que son robustas a la categoría base seleccionada.

Respecto a la sobreeducación, se debe recordar en primer lugar que la brecha entre hombres y mujeres se invirtió en el periodo analizado. Así, mientras en 1995 la sobreeducación en los varones superaba en 5 puntos a la existente en las mujeres, en 2006 el gap en sobreeducación era superior en las mujeres en casi 3 puntos. La primera cuestión que conviene resaltar es que esta evolución de la divergencia entre hombres y mujeres es atribuible tanto al aumento (en términos absolutos) de la contribución de la diferencia en el impacto de las características, que prácticamente se cuadruplicó entre 1995 y 2006, como al descenso de la contribución de las diferencias en su dotación. Así, la mayor probabilidad de sobreeducación en los hombres en 1995 tuvo su origen en una dotación de características que los hacía más propensos a ese tipo de desajuste que las mujeres. En 2006, la inversión de la diferencia tuvo que ver en parte con la sustancial disminución de ese efecto y con el notable incremento de la diferencia en rendimientos, desfavorable a las mujeres. En cualquier caso, debe destacarse que después de controlar por las diferencias en características, el diferencial en la probabilidad es desfavorable para las mujeres, por lo que se puede concluir que la tasa de sobreeducación neta es claramente mayor en ellas, particularmente en el último año analizado.

Además la descomposición detallada nos permite vislumbrar un hecho adicional, la interesante evolución experimentada por el entorno laboral en el gap en sobreeducación por género. Observamos como en 1995 la contribución total (por dotación y rendimientos) de las divergencias derivadas de los puestos de trabajo y de las empresas habría disparado la brecha por género hasta alcanzar 22 puntos porcentuales a favor del colectivo varón. Esta importancia del entorno laboral se debe básicamente a dos factores: diferencias en las características de los puestos de trabajo imputable a una importante segregación ocupacional (13.79 puntos) y, aunque en un segundo plano, a las diferencias en los efectos de las características empresariales, básicamente asociadas a la propiedad y al sector de actividad. Sin embargo, y aún manteniendo cierta relevancia, la contribución de estos factores del entorno laboral se redujo de forma importante en 2006. Nótese que si en ese año hombres y mujeres únicamente hubiesen divergido en estos aspectos, la proporción de varones sobreeducados superaría en 9.5 puntos a la existente para las mujeres, frente a los casi 14 puntos de diferencia originados por ese tipo de características una década antes. Esta evolución de la contribución de las características del entorno laboral es debida a que en 2006 hombres y mujeres presentaban prácticamente idéntico impacto del conjunto de características

asociadas al puesto de trabajo y a la empresa y al desvanecimiento de la contribución de las diferencias en las ocupaciones, asociada a la progresiva homogeneización en las ocupaciones desempeñadas por ambos colectivos. La única característica que evolucionó en sentido contrario fue la del sector de actividad, dado que la diferente distribución en ramas de actividad de hombres y mujeres pasó de tener una contribución de 1.3 puntos en 1995 a casi 6 en 2006, siempre propiciando una mayor tasa de sobreeducación en los hombres.

El efecto de las diferencias en dotación y rendimiento del capital humano, imputable casi en su totalidad a la educación, no es menos interesante que la presentada por el entorno laboral. Así, la contribución de la diferencia en la dotación educativa, prácticamente estable a lo largo del periodo, hubiera motivado una mayor sobreeducación en las mujeres. Pero, por el contrario, la diferencia de género en el efecto asociado a la educación, aún reduciéndose en 5 puntos entre 1995 y 2006, habría ocasionado que la probabilidad media de sobreeducación en los hombres superase en 16 puntos porcentuales la de las mujeres al final del periodo.

La evolución de las tasas de infraeducación, la diferencia por género en las mismas y el origen del gap no resulta menos interesante. Durante el periodo 1995-2006 el aumento de la infraeducación fue un fenómeno observable para el conjunto de asalariados, pero especialmente intenso en el colectivo masculino (Tabla 1). En 2006 un tercio de los trabajadores varones estaban infraeducados, 4.6 puntos más que las mujeres. De hecho, el sustancial aumento en la incidencia de la infraeducación en los hombres provocó la inversión en la brecha de género, ya que en 1995 eran las mujeres las más afectadas por este tipo de desajuste educativo (3.6 puntos más que el colectivo masculino). Este cambio, como se desprende de la Tabla 5, se origina tanto por la evolución de las diferencias en características como por la de sus efectos. En líneas generales, podemos decir que una cierta homogeneización en la distribución de características entre hombres y mujeres (mayor igualdad en dotación de capital humano, en puestos de trabajo desempeñados y en el perfil de las empresas que les contrata) provocó una disminución en la contribución de la diferencia en dotaciones. Este hecho contrasta con un aumento en la contribución del impacto de las características observables en las diferencias de género en la probabilidad de infraeducación.

Un análisis más detallado muestra cómo los aspectos asociados con el puesto de trabajo y la empresa han desempeñado un papel relevante en los dos años considerados, aunque sus elementos han mostrado una evolución dispar. Así, si hombres y mujeres sólo hubiesen divergido en su entorno laboral, la brecha en infraeducación, mayor en el colectivo femenino, se hubiera situado en torno a los nueve puntos, tanto en 1995 como en 2006. Y, de igual forma que en el caso de la

sobreeducación, fue la segregación ocupacional el factor explicativo más relevante al inicio del periodo. Su magnitud es tal que, de no existir diferencias en otros factores, habría ocasionado que el diferencial por sexo en 1995 alcanzase casi 11 puntos porcentuales. Esta contribución de la diferencia en el tipo de puestos de trabajo desempeñados por hombres y mujeres es mitigada, en parte, por la de las diferencias vinculadas con el capital humano, tanto en dotación como en rendimiento. Y son fundamentalmente las diferencias en el nivel educativo las que habrían ocasionado que la proporción de hombres infraeducados hubiese superado en más de 15 puntos porcentuales a la presentada por el colectivo femenino. De ellos, seis puntos corresponderían al mayor nivel de estudios de las mujeres y 9.5 puntos por las divergencias en el impacto de la educación en la probabilidad de infraeducación.

El escenario que describen los resultados detallados para 2006 es completamente distinto. Respecto al entorno profesional (puesto de trabajo y empresa), aunque su peso en conjunto parece mantenerse estable no puede decirse lo mismo respecto a la contribución de sus componentes. Destaca, en primer lugar, la pérdida de importancia de la segregación ocupacional en la explicación del gap en infraeducación entre hombres y mujeres (al final del periodo su contribución no representa ni una cuarta parte de la alcanzada en 1995). Sin embargo, aumenta la contribución de las divergencias en el impacto de otras características empresariales como el sector de actividad y, muy especialmente, la participación pública en la empresa. La diferencia en el efecto asociado al tipo de propiedad y control de la empresa⁶ pasa de tener una contribución casi testimonial en 1995 (representando poco más de un punto) a ocasionar una brecha superior a ocho puntos, por una mayor incidencia de la infraeducación en las mujeres. Pero sin duda, en ese último año analizado, es la contribución del capital humano al diferencial en la probabilidad de infraeducación la más contundente cuantitativamente. Y ello no es imputable a las diferencias en su dotación, que se concentran básicamente en educación y que redujeron a la mitad su contribución al diferencial entre 1995 y 2006. Fue fundamentalmente la discrepancia en el impacto de la educación, que aumentó en términos absolutos su contribución y que, además, invirtió el sentido de su efecto. Tal es la repercusión de este aspecto que si en 2006 hombres y mujeres sólo se hubiesen diferenciado en el impacto que tuvo el nivel de estudios en la probabilidad de estar infraeducado, la incidencia de la infraeducación en el colectivo femenino habría superado en 15.6 puntos a la del masculino.

Por último, es justo destacar que si finalmente las mujeres no mostraron mayor proporción de infraeducación que los varones en el último año analizado, a pesar de favorecerlo tanto las

⁶ Cabe recordar que por las características de la muestra de la EES sólo podemos distinguir entre empresas con control totalmente privado y empresas con participación pública, ya que en la misma no se incluyen trabajadores empleados en el sector público.

características empresariales y del puesto de trabajo como los efectos asociados al capital humano, es porque la parte atribuible a la constante compensó estos efectos. En este sentido, cabe indicar que dado que la descomposición se obtiene a través de la estimación del modelo probabilístico que incorpora las restricciones de identificación propuestas en Gardeazabal y Ugidos (2004), las diferencias correspondientes a las constantes son atribuibles a características comunes a, por una parte, todas las mujeres y, por otra, todos los hombres en la muestra. Por tanto, las diferencias en ese tipo de características común en el seno de ambos colectivos habría cobrado una destacada importancia para más que compensar las diferencias atribuibles a las características observadas consideradas en este trabajo.

5. Reflexiones finales

Entre 1995 y 2006 el porcentaje de trabajadores sobreeducados en España se mantuvo relativamente estable en valores próximos a un tercio. Por su parte, la tasa de infraeducación experimentó un notable aumento a lo largo de ese periodo, de forma que el porcentaje de trabajadores adecuadamente educados en la economía española al final del periodo analizado se encontraba por debajo del 40%, valor ciertamente preocupante incluso tras tener en cuenta las críticas que se pueden formular al criterio estadístico utilizado en la definición de la adecuación educativa en este trabajo. Adicionalmente, hemos mostrado como la diferencia de género tanto en sobreeducación como en infraeducación es sustancial y no constante a lo largo del tiempo.

Con el objetivo de indagar en el origen de esas diferencias de género, la hipótesis de este trabajo ha sido la de suponer que la mayor sobreeducación y la menor infraeducación netas en las mujeres tiene su origen, no únicamente en diferencias en la dotación y en el impacto de características personales sino también en las relativas a los puestos de trabajo y a los establecimientos en los que prestan sus servicios trabajadores y trabajadoras. Los resultados obtenidos han confirmado esta hipótesis, revelando además que la contribución de ese tipo de características fue más intensa al inicio que al final del periodo analizado. Así, por ejemplo, su efecto es determinante para explicar las diferencias de género en la tasa de sobreeducación en 1995, básicamente por las diferencias entre hombres y mujeres en la distribución en el tipo de ocupaciones. La homogeneización en éstas en 2006 está detrás de la fuerte reducción atribuible a esta componente.

Por otra parte, el aumento en el diferencial en la tasa de infraeducación, originado principalmente por el importante incremento en su incidencia en los hombres, junto a la evidencia obtenida en la

descomposición que sugiere, también en este caso, un destacado papel de la contribución de las diferencias en características del puesto de trabajo y del establecimiento, confirman la relevancia de considerar también esta vertiente del desajuste educativo. De hecho, las cifras correspondientes a 2006 indican que, cualitativamente y desde la visión de género, la infraeducación tiene un papel superior al de la sobreeducación, por lo que parece necesario plantearse la incidencia particular de tal fenómeno en, por ejemplo, la brecha salarial entre hombres y mujeres.

En todo caso, conviene resaltar que nuestro interés en el papel desempeñado por las características empresariales ha motivado la utilización de la información contenida en las muestras de la EES para España. Y, en consecuencia, la no consideración de factores como los familiares y los geográficos, así como la posible existencia de selección muestral, que pueden estar afectando a las probabilidades de inadecuación educativa, por la ausencia de ese tipo de información en la EES. En este sentido, es posible que la diferencia de probabilidades atribuible en algunos casos a los términos independientes estimados tengan, al menos en parte que ver con las diferencias por las que no estamos controlando en nuestros resultados. En todo caso, nótese que esta cuestión no invalida en términos generales la evidencia presentada respecto a nuestra hipótesis acerca de la contribución de las características empresariales y laborales.

Por último, incidir en el hecho de que el análisis de descomposición se ha basado en la estimación de un modelo de probabilidad binomial, considerando por tanto de forma separada los fenómenos de la sobre y la infraeducación. A pesar de que la estimación de los efectos marginales de la mayoría de características observadas era consistente con la estimada a partir de un modelo logit multinomial que considera conjuntamente ambos fenómenos, la obtención de la descomposición detallada basada en ese tipo de modelo multinomial se encuentra entre el listado de prioridades de nuestra investigación futura.

6. Bibliografía

Aguilar, M. I.; Navarro, M. L. (2003), "La incidencia del desajuste educativo en el primer empleo de los jóvenes", XII Jornadas de la Asociación de Economía de la Educación (AEDE), Madrid.

Alba-Ramírez, A. (1993), "Mismatch in The Spanish Labor Market: Overeducation?", *Journal of Human Resources*, vol. 28, pp. 259-278.

Amuedo-Dorantes, C.; De la Rica, S. (2006), "The Role of Segregation and Pay Structure on the Gender Wage Gap: Evidence from Matched Employer-Employee Data for Spain", *Contributions to Economic Analysis & Policy*, 5 (1), pp. 1-32.

Beneito, P., Ferri, J., Moltó, M. L.; 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.; Battu, H. (2003), "The Theory of Differential Overqualification: Does It Work?", *Scottish Journal of Political Economy*, 50 (1), pp. 1-16.
- De la Rica, S. Dolado, J., Llorens, V. (2008), "Ceilings or floors? Gender wage gaps by education in Spain", *Journal of Population Economics*, 21(3), pp. 751-776.
- De la Rica, S.; Ugidos, A. (1995), "¿Son las diferencias en capital humano determinantes en las diferencias salariales observadas entre hombres y mujeres?", *Investigaciones Económicas*, XIX(3), pp. 395-414.
- Duncan, G.; Hoffman, S. (1981), "The Economic Value of Surplus Education", *Economics of Education Review*, vol. 1, n. 1, pp.75-86.
- Frank, R. (1978), "Why Women Earn Less: The Theory and Estimation of Differential Overqualification", *American Economic Review*, vol. 68, n. 3, pp. 360-373.
- Freeman, R. B. (1976). *The Overeducated American*. London: Academic Press.
- García Montalvo, J. (1995), *Empleo y Sobrecualificación: El Caso Español*, Documento de Trabajo FEDEA 95-20.
- García-Serrano, C. y Malo, M. A. (1996), "Desajuste Educativo y Movilidad Laboral en España", *Revista de Economía Aplicada*, 4 (11), pp. 105-131.
- García-Serrano, C.; Malo, M. A. (1997), "¿Es Diferente el Desajuste Educativo de las Mujeres?", *Información Comercial Española. Revista de Economía*, 760, pp. 117-128.
- Groot, W.; Maassen Van den Brink, H. (2000), "Overeducation in the Labor Market: A Meta-Analysis", *Economics of Education Review* vol. 19, pp. 149-158.
- Hartog, J.; Oosterbeek, H. (1988), "Education, Allocation and Earnings in the Netherlands: Overschooling?", *Economics of Education Review* vol. 7, n. 2, pp.185-194.
- Johansson, M.; Katz; K. (2007), *Wage differences between women and men in Sweden – the impact of skill mismatch*, Institute for Labour Market Policy Evaluation (IFAU) Report 2007:13.
- Karakaya, G., Plasman, R.; Rycx, F. (2007), "Overeducation on the Belgian labour market: evaluation and analysis of the explanatory factors through two types of approaches", *Compare: A Journal of Comparative and International Education*, 37 (4), pp. 513-532.
- Kiker, B.; Santos, M. (1991), "Human Capital and Earnings in Portugal", *Economics of Education Review*, vol. 10, n. 3, pp. 187-203.
- Kiker, B., Santos, M.; Mendes de Oliveira, M. (1997), "Overeducation and Undereducation: Evidence for Portugal", *Economics of Education Review*, vol. 16, n. 2, pp. 111-125.
- Marzo-Navarro, M. (2007), "The educational gap in higher education: the Spanish case", *Journal of Education and Work*, 20 (2), pp. 123-137
- Mcgoldrick, K.; 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.; Kiker, B. (2000), "The Role of Human Capital and Technological Change in Overeducation", *Economics of Education Review*, vol. 19, pp. 199-206.
- Ng, Y, C. (2003), "Earnings and "skill" allocation in the Canadian labor market", *International Journal of Manpower*, 24 (8), pp. 964-980.
- Oliver, J.; 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 14.
- Robst, J. (2007), "Education, College Major, and Job Match: Gender Differences in Reasons for Mismatch", *Education Economics*, 15 (2), pp. 159-175

Rumberger, R. (1987). "The Impact of Surplus Schooling on Productivity and Earnings", *Journal of Human Resources*, vol. XXII, n. 1, pp. 24-50.

Sanromá, E.; Ramos, R. (2004), " Sobreeducación y mercados de trabajo locales en España", VII Encuentro de Economía Aplicada, Vigo.

Sicherman, N. (1991), "Overeducation in the Labour Market", *Journal of Labor Economics* vol. 9, pp. 101-122.

Simón, H.; Ramos, R.; Sanromá, E. (2008), "La evolución de las diferencias salariales por razón de sexo en España (1995-2002)", *Revista de Economía Aplicada*, 48 (XVI), pp. 37-68.

Verdugo, R.; Verdugo, N. (1989), "The Impact of Surplus Schooling on Earnings", *Journal of Human Resources*, vol. 24, n. 4, pp. 629-643.

Yun, M. (2004), "Decomposing differences in the first moment", *Economics Letters*, Elsevier, vol. 82(2), pp. 275-280.

7. Tablas y gráficos

Tabla 1. Incidencia del desajuste educativo y diferencia por género

<i>Porcentaje de trabajadores</i>	Todos		Hombre		Mujer	
	1995	2006	1995	2006	1995	2006
Sobreeducado	31,98%	29,93%	33,18%	29,02%	28,1%	31,83%
Adecuadamente educado	48,03%	38,04%	47,67%	37,47%	49,16%	39,23%
Infraeducado	19,99%	32,03%	19,14%	33,51%	22,74%	28,93%
<i>Diferencia Hombre-Mujer</i>	1995	2006				
Sobreeducado	5,08%	-2,81%				
Infraeducado	-3,6%	4,58%				

Tabla 2. Probabilidad de sobreeducación

<i>Efectos marginales</i>		Todos		Hombres		Mujeres	
		1995	2006	1995	2006	1995	2006
Capital Humano	Hombre	-0,0145***	-0,0314***				
	Educación	0,171***	0,0830***	0,180***	0,0581***	0,129***	0,124***
	Experiencia	0,00186***	0,00150***	0,00360***	0,00103***	-0,00161***	0,0003
	Experiencia ²	-6,89e-05***	-2,18e-05***	-0,000108***	-1,87e-05***	1,22E-05	2,39e-05*
	Antigüedad	-0,000816	-0,000186	-0,00224***	-0,000465***	0,00547***	0,000693
	Antigüedad ²	-1,74E-05	-1,26e-05**	2,91E-05	3,11E-06	-0,000226***	-5,18e-05***
Caract. del puesto de trabajo	Jornada completa	-0,0682***	-0,0581***	-0,000677	-0,0136***	-0,0701***	-0,0870***
	Contrato indefinido	-0,0398***	-0,0120***	-0,0409***	-0,0142***	-0,0350***	-0,0052
	Directivos	-0,191***	-0,0712***	-0,223***	-0,0481***	-0,0781***	-0,116***
	Titulados universitarios	-0,225***	-0,0905***	-0,249***	-0,0574***	-0,127***	-0,161***
	Técnicos y administrativos	-0,476***	-0,195***	-0,415***	-0,109***	-0,737***	-0,382***
	Trabajadores de servicios	-0,138***	-0,0495***	-0,127***	-0,0223***	-0,103***	-0,111***
	Trabajadores cualificados	0,101***	-0,0517***	0,116***	-0,0339***	0,0604***	-0,0714***
	Operadores de maquinaria	0,301***	-0,0664***	0,293***	-0,0402***	0,411***	-0,117***
Características de la empresa	Industrias extractivas	0,145***	0,105***	0,246***	0,143***	-0,0351***	0,161***
	Industria manufacturera	0,0635***	0,0621***	0,135***	0,0550***	-0,0609***	0,0842***
	Construcción	0,0468***	0,443***	0,121***	0,493***	-0,0420***	0,247***
	Comercio y Hostelería	0,0686***	0,0664***	0,137***	0,0994***	-0,0469***	0,0459***
	Transporte y comunicaciones	-0,123***	0,107***	-0,116***	0,128***	-0,0694***	0,130***
	Número de trabajadores	-2,55e-05***	-7,95E-07	-3,53e-05***	6,49E-08	-7,00E-06	-1,68E-06
	Empresa pública	0,139***	0,0756***	0,194***	0,0454***	-0,00539	0,147***
	Mercado local	-0,0235***	-0,00176	-0,0264***	-0,00612***	-0,00249	0,0213***
	Convenio sectorial	0,0577***	-0,0207***	0,0712***	-0,00802***	0,0288***	-0,0446***
	Convenio provincial	0,0867***	-0,00677***	0,0903***	-0,00650***	0,0566***	0,00202
Número de observaciones	132684	136261	101360	92316	31324	43945	
Pseudo R ²	0,5802	0,6284	0,5744	0,6692	0,6487	0,5736	

Los modelos estimados también incluyen efectos fijos regionales.

***, **, * indican efectos marginales estadísticamente significativos al 1%, al 5% y al 10%, respectivamente.

Tabla 3. Probabilidad de infraeducación

<i>Efectos marginales</i>		Todos		Hombres		Mujeres	
		1995	2006	1995	2006	1995	2006
Capital Humano	Hombre	0,0108***	0,0713***				
	Educación	-0,0697***	-0,158***	-0,0654***	-0,170***	-0,0651***	-0,135***
	Experiencia	0,00260***	-0,00183***	0,00129***	-0,00139**	0,00348***	0,000735
	Experiencia ²	-5,10e-05***	0,00000838	-2,97e-05***	0,0000107	-5,33e-05***	-5,16e-05***
	Antigüedad	-0,000420*	-0,000908**	-0,000277	0,0000152	-0,00159***	-0,00186***
	Antigüedad ²	2,81e-05***	7,37e-05***	2,30e-05***	4,41e-05***	6,12e-05***	8,50e-05***
Caract. del puesto de trabajo	Jornada completa	0,0303***	0,0874***	0,0078	0,0453***	0,0273***	0,0906***
	Contrato indefinido	0,0155***	0,0162***	0,0156***	0,0362***	0,0105***	0,00336
	Directivos	0,968***	0,907***	0,970***	0,908***	0,974***	0,906***
	Titulados universitarios	0,973***	0,921***	0,973***	0,914***	0,984***	0,933***
	Técnicos y administrativos	0,688***	0,936***	0,691***	0,939***	0,664***	0,849***
	Trabajadores de servicios	0,107***	0,431***	0,0527***	0,265***	0,269***	0,454***
	Trabajadores cualificados	-0,0421***	0,245***	-0,0473***	0,212***	-0,0143***	0,196***
	Operadores de maquinaria	-0,103***	0,343***	-0,106***	0,279***	-0,0692***	0,428***
Características de la empresa	Industrias extractivas	-0,0307***	-0,0650***	-0,0408***	-0,0936***	0,0739***	-0,033
	Industria manufacturera	-0,0153***	-0,0854***	-0,0457***	-0,135***	0,0592***	-0,0569***
	Construcción	-0,0120***	-0,165***	-0,0320***	-0,202***	0,108***	-0,104***
	Comercio y Hostelería	-0,0248***	-0,0567***	-0,0372***	-0,0876***	0,0985***	-0,0334***
	Transporte y comunicaciones	-0,0025	-0,0637***	-0,0182***	-0,0853***	0,0192**	-0,0639***
	Número de trabajadores	-1,03e-05***	-4,54e-06**	-8,07e-06***	-2,29E-06	-9,09e-06***	-7,11e-06***
	Empresa pública	-0,0133***	-0,0902***	-0,0220***	-0,0718***	0,0351**	-0,105***
	Mercado local	-0,00624***	-0,00931***	-0,000978	-0,000235	-0,0284***	-0,0277***
	Convenio sectorial	-0,000985	0,0183***	-0,0102***	-0,00858*	0,00323	0,0508***
	Convenio provincial	-0,00939***	-0,0187***	-0,00633***	-0,0143***	-0,0173***	-0,0155**
Número de observaciones	132684	136261	101360	92316	31324	43945	
Pseudo R ²	0,588	0,6761	0,5854	0,7084	0,6658	0,6247	

Los modelos estimados también incluyen efectos fijos regionales.

***, **, * indican efectos marginales estadísticamente significativos al 1%, al 5% y al 10%, respectivamente.

Tabla 4. Descomposición detallada de la diferencia de género en la probabilidad de sobreeducación

	1995		2006	
	Diferencias en Características	Diferencias en Efectos	Diferencias en Características	Diferencias en Efectos
Diferencia por género	0,0508		-0,0281	
Total	0,0642	-0,0134	0,0223	-0,0504
Capital Humano	-0,0845	0,2098	-0,0764	0,1588
Educación	-0,0753	0,2198	-0,0748	0,1570
Experiencia	-0,0065	-0,0265	0,0012	0,0039
Antigüedad	-0,0027	0,0165	-0,0028	-0,0022
Características Puesto de Trabajo	0,1379	0,0059	0,0364	0,0079
Jornada	0,0000	-0,0167	-0,0078	0,0041
Tipo de Contrato	-0,0018	-0,0016	0,0001	-0,0029
Ocupación	0,1397	0,0242	0,0441	0,0068
Características de la Empresa	0,0108	0,0615	0,0623	-0,0112
Sector	0,0129	0,0235	0,0592	-0,0086
Tamaño	0,0004	0,0018	0,0000	0,0002
Propiedad	0,0003	0,0327	0,0002	0,0043
Mercado	-0,0002	0,0034	0,0005	-0,0054
Convenio	-0,0022	-0,0005	0,0016	-0,0002
Región	-0,0003	0,0006	0,0009	-0,0015
Constante	-	-0,2906	-	-0,2059

Tabla 5. Descomposición detallada de la diferencia de género en la probabilidad de infraeducación

	1995		2006	
	Diferencias en Características	Diferencias en Efectos	Diferencias en Características	Diferencias en Efectos
Diferencia por género	-0,0360		0,0458	
Total	-0,0587	0,0227	-0,0157	0,0615
Capital Humano	0,0597	0,0786	0,0284	-0,1559
Educación	0,0591	0,0949	0,0281	-0,1562
Experiencia	-0,0011	-0,0194	-0,0010	-0,0066
Antigüedad	0,0017	0,0030	0,0013	0,0070
Características Puesto de Trabajo	-0,1035	0,0001	-0,0207	-0,0096
Jornada	0,0012	-0,0080	0,0045	-0,0166
Tipo de Contrato	0,0016	0,0001	0,0000	0,0082
Ocupación	-0,1063	0,0079	-0,0252	-0,0012
Características de la Empresa	-0,0149	0,0303	-0,0234	-0,0415
Sector	-0,0152	0,0116	-0,0236	0,0263
Tamaño	0,0002	0,0003	0,0001	0,0019
Propiedad	-0,0001	0,0129	-0,0001	-0,0825
Mercado	0,0000	0,0049	0,0000	0,0103
Convenio	0,0013	-0,0002	0,0001	-0,0053
Región	-0,0011	0,0008	0,0001	0,0078
Constante	-	-0,0863	-	0,2685

Anexo 1. Descriptivos de las principales variables consideradas en el análisis

	Todos		Hombres		Mujeres	
	1995	2006	1995	2006	1995	2006
Hombre	76,4%	67,7%	100,0%	100,0%	0,0%	0,0%
Mujer	23,6%	32,3%	0,0%	0,0%	100,0%	100,0%
Años de estudio	9,0	10,3	8,9	10,2	9,4	10,6
Años de experiencia potencial	23,3	22,2	24,7	23,0	19,1	20,3
Años de antigüedad	10,3	7,7	10,9	8,5	8,3	6,1
Jornada completa	96,2%	88,7%	98,3%	95,6%	89,3%	74,2%
Jornada parcial*	3,8%	11,3%	1,7%	4,4%	10,7%	25,8%
Contrato indefinido	73,3%	74,8%	74,7%	74,8%	68,7%	75,0%
Contrato temporal*	26,7%	25,2%	25,3%	25,2%	31,3%	25,0%
Directivos	4,0%	2,8%	4,8%	3,4%	1,3%	1,6%
Titulados universitarios	4,6%	5,0%	4,8%	5,0%	3,8%	5,1%
Técnicos y administrativos	25,4%	25,6%	20,5%	20,3%	41,5%	36,8%
Trabajadores de servicios	7,9%	8,0%	6,3%	4,6%	12,9%	15,3%
Trabajadores cualificados	21,1%	22,5%	24,8%	29,5%	8,9%	7,8%
Operadores de maquinaria	24,4%	21,1%	27,0%	25,4%	15,9%	11,9%
No calificados*	12,7%	15,0%	11,7%	11,9%	15,8%	21,5%
Industrias extractivas	3,5%	0,9%	4,0%	1,1%	1,7%	0,2%
Industria manufacturera	44,5%	52,3%	45,2%	57,7%	42,4%	40,9%
Construcción	9,3%	9,7%	10,3%	13,2%	5,9%	2,4%
Comercio y Hostelería	7,4%	15,8%	9,1%	11,1%	1,8%	25,5%
Transporte y comunicaciones	5,6%	4,8%	6,2%	5,5%	4,0%	3,4%
Intermediación financiera y act. inmobiliarias*	29,7%	16,6%	25,2%	11,4%	44,2%	27,6%
Número de trabajadores	177,2	265,8	174,1	235,3	187,3	329,9
Empresas con participación pública	2,1%	0,6%	2,1%	0,6%	1,9%	0,4%
Empresas privadas*	97,9%	99,4%	97,9%	99,4%	98,1%	99,6%
Mercado local	88,7%	84,4%	88,9%	83,5%	87,8%	86,3%
Otros mercados*	11,3%	15,6%	11,1%	16,5%	12,2%	13,7%
Convenio sectorial	34,6%	37,3%	32,1%	34,8%	42,7%	42,5%
Convenio provincial	42,2%	47,0%	43,3%	47,7%	38,5%	45,5%
Convenio de empresa*	54,0%	47,2%	56,8%	48,8%	45,1%	43,8%
Número de observaciones	132684	136261	101360	92316	31324	43945

* Categoría de referencia en los modelos probit