

La sobre-educación en Cataluña y su relación con el conocimiento del catalán¹

Maite Blázquez

*Universidad Autónoma de Madrid
maite.blazquez@uam.es*

Silvio Rendon

*Instituto Tecnológico Autónomo de México
srendon@itam.mx*

RESUMEN

En economías como Cataluña caracterizadas por la coexistencia de varias lenguas los individuos que no dominen una lengua podrían intentar compensar esta deficiencia con un mayor nivel de educación formal, generándose así una relación entre la falta de cualificaciones lingüísticas y la incidencia del fenómeno de la sobre-educación. Un análisis descriptivo de los datos de los Institutos Catalán y Nacional de Estadística (IDESCAT, INE) sugiere que aquellos individuos con mayor nivel de conocimiento del catalán tienen un mayor riesgo de estar sobre-educados. Sin embargo,

¹ Agradecemos a los participantes en la XX Conferencia Anual de la European Society for Population Economics en Verona, de seminarios en la U. Pompeu Fabra, U. de Girona, FEDEA, U. Autónoma of Madrid, y U. of Salamanca, así como a Núria Quella, Miguel Sánchez Romero y a un evaluador anónimo por sus comentarios y sugerencias. Cualquier error u omisión es exclusiva responsabilidad de los autores.

cuando controlamos por diversas variables, las estimaciones conducen al resultado contrario: saber catalán tiene un efecto negativo sobre la probabilidad de estar sobre-educado. No obstante, este efecto aunque significativo a nivel individual y robusto al carácter endógeno del conocimiento del catalán, no resulta significativo a nivel agregado.

Clasificación JEL: J61, J70, J31, I20.

Palabras Clave: Sobre-educación, Cualificaciones Lingüísticas, Inmigración.

1. INTRODUCCIÓN

En un mercado laboral competitivo, los individuos tienden a compensar la falta de algunas cualificaciones adquiriendo más de otras. En este sentido, la escasez de cualificaciones lingüísticas puede llevar a los individuos a aceptar empleos para los que están sobre-educados, es decir, para los que tienen cualificaciones formales en demasía. Dado que la coexistencia de varias lenguas suele ser frecuente en la mayoría de países, este tipo de desajuste educativo puede darse no sólo entre los inmigrantes sino también entre los nativos. Este puede ser el caso de países o regiones que han experimentado algún cambio en materia lingüística, como por ejemplo el hecho de que una lengua de uso familiar o popular pasa a ser lengua oficial.

En la literatura, el fenómeno de sobre-educación ha tratado de explicarse como una compensación por la falta de alguna forma de capital humano, como experiencia, habilidad o formación en el puesto de trabajo (Sicherman 1991, Alba-Ramírez 1993, Groot 1993, 1996, Groot y Maasen van den Brink 1996, García Serrano y Malo, 1996). En concreto, los resultados del trabajo de García-Serrano y Malo (1996) sugieren la existencia de un cierto grado de sustitución entre formación general y formación específica. En este sentido, y dado que las cualificaciones lingüísticas constituyen una forma específica de capital humano la sobre-educación puede surgir, por tanto, como consecuencia de una carencia de este tipo de habilidades.

Encontramos numerosos los trabajos en la literatura que han proporcionado evidencia sobre cómo el conocimiento de la lengua predominante en una economía permite mejorar de manera significativa la situación laboral de un individuo, especialmente en términos de ganancias, probabilidades de empleo, y riesgo de desempleo (Niesing et al. 1994, Bratsberg & Ragan 1998, Borjas 1999, Schaafsma & Sweetman, 2001, Chiswick & Miller 2002; 2003, Rendon, 2006) . Sin embargo, los pocos trabajos que relacionan desajuste educativo y cualificaciones lingüísticas no proporcionan una evidencia clara de la relación existente entre ambas. Además, estos trabajos

se centran en un contexto de asimilación de inmigrantes en una economía doméstica y no en el caso sociedades multilingües.

En un extremo de esta escasa literatura encontramos evidencia de que el dominio de la lengua predominante aumenta la probabilidad de sobre-educación. Así, el trabajo de Battu y Sloane (2002) analiza el problema de la sobre-educación entre las minorías étnicas en Gran Bretaña y encuentra que el dominio del inglés aumenta tanto las probabilidades de sobre-educación como infra-educación. En esta misma línea, y para el caso de Australia, el trabajo de Linsley (2005) encuentra que los inmigrantes hombres procedentes de países de habla no inglesa presentan unas tasas de sobre-educación significativamente mayores que los nativos.

Otros trabajos, por el contrario, sugieren que las cualificaciones lingüísticas aumentan el riesgo de sobre-educación. Así, el trabajo de Green et al., (2004) analiza la incidencia y el impacto de la sobre-educación entre los inmigrantes recién llegados a Australia. Sin embargo, a diferencia del trabajo de Linsley, estos autores encuentran que los inmigrantes procedentes de países de habla no inglesa, especialmente asiáticos, presentan unas mayores tasas de sobre-educación debido, en parte, a un escaso conocimiento de la lengua inglesa. Un resultado similar se obtiene en el trabajo de Barrett et al., (2005) que analiza el impacto de la inmigración sobre el mercado laboral irlandés. Estos autores encuentran que los inmigrantes constituyen un grupo altamente cualificado pero que, por lo general, suelen ocupar empleos que no se ajustan íntegramente a su nivel de formación, salvo en el caso de inmigrantes procedentes de Estados Unidos o Reino Unido que no experimentan diferencias significativas con respecto a los nativos, lo que hace sospechar que este desajuste educativo puede estar relacionado con las habilidades en el conocimiento de la lengua inglesa, de manera que aquellos inmigrantes procedentes de países de habla no inglesa experimentan un mayor riesgo de estar sobre-educados.

El presente artículo pretende constituir una nueva aportación a la literatura de sobre-educación, tratando de medir el efecto del conocimiento de una lengua sobre la probabilidad de estar sobre-educado en un mercado laboral donde coexisten varias lenguas, y donde las cualificaciones lingüísticas juegan un papel importante a la hora de explicar la situación laboral de los trabajadores. Como en otras economías caracterizadas por la diversidad lingüística, en Cataluña la intervención de los organismos públicos a través de políticas activas en materia lingüística ha tenido implicaciones económicas directas. En particular, esta política contribuyó a aumentar el valor económico del catalán ya que aquellos individuos que saben catalán tienen unas mayores probabilidades de empleo (Rendon 2006). Nuestro trabajo trata de dar un paso más en esta línea al tratar de analizar el efecto de saber catalán sobre la

probabilidad de que un individuo encuentre un trabajo que se ajuste a su nivel de cualificación. Dado que los cambios introducidos por la política en materia lingüística afectan tanto a «inmigrantes» como a «nativos», cabría esperar una mayor incidencia del fenómeno de sobre-educación entre aquellos individuos con un bajo nivel de conocimiento del catalán. No obstante, los resultados obtenidos a partir de datos de los censos de 1991 y 1996 nos llevan a rechazar la hipótesis de «sustitución de cualificaciones», puesto que una vez que estimamos controlando por varios factores y tenemos en cuenta el carácter endógeno del conocimiento del catalán, el efecto medio de este último sobre la probabilidad de estar sobre-educado no resulta significativo.

El resto del artículo se organiza de la forma siguiente. En la próxima sección nos centramos en el fenómeno de sobre-educación. La Sección 3 proporciona una revisión sobre la política en materia lingüística llevada a cabo en Cataluña desde los años cuarenta hasta la actualidad. En la Sección 4 describimos la base de datos y discutimos los principales resultados de la estadística descriptiva. La Sección 5 recoge los resultados de las estimaciones mientras que en la Sección 6 se presentan las principales conclusiones.

2. LA SOBRE-EDUCACIÓN COMO DESAJUSTE EDUCATIVO

En los últimos años, muchos países han realizado un gran esfuerzo de inversión en educación que se ha traducido en que un elevado porcentaje de la población potencialmente activa ha alcanzado un alto nivel educativo. Esta expansión educativa, a su vez, ha ido acompañada por un aumento en la oferta de trabajos cualificados, así como por un mayor requisito de habilidades para desempeñar dichos trabajos. No obstante, este aumento en la demanda de trabajadores con un alto nivel educativo no ha sido capaz de absorber el aumento en la oferta de los mismos, lo que ha dado lugar a la aparición del fenómeno de sobre-educación, y a una disminución en la demanda de trabajadores con bajo nivel de cualificación.

Como señalan Wieling y Borghans (2001), la sobre-educación es el principal mecanismo de ajuste de un mercado laboral caracterizado por exceso de oferta de trabajadores altamente cualificados. En estas situaciones, las empresas sustituyen trabajadores con bajo nivel de cualificación por trabajadores más cualificados, y así estar provistos de este tipo de mano de obra cuando es escasa (Dupuy y de Grip 2002). Además, el aumento en la oferta de trabajadores altamente cualificados puede reforzar y acelerar el cambio tecnológico sesgado, ya que las empresas pueden comple-

mentar este tipo de fuerza laboral con tecnologías intensivas en factor capital (Acemoglu 2002).

Este concepto de sobre-educación fue introducido por Thurow (1975) y Freeman (1976) en el contexto de la economía americana en los años 70². Encontramos una vertiente de la literatura que considera la sobre-educación como un fenómeno temporal: bien como la transición hacia una nueva situación de equilibrio (Becker, 1957), o como parte de la carrera profesional de los trabajadores (Rosen 1972, Jovanovic 1979, Sicherman y Galor 1990, Sicherman 1991, Robst 1995a, Groot 1993,1996, Groot y Maasen van den Brink 1997, Dekker, de Grip y Heijke 2002)³. Sin embargo, otros enfoques consideran la sobre-educación como un fenómeno de carácter más permanente. Es el caso de la Teoría de Señalización (Spence 1973) y la Teoría de Competencia por los puestos de trabajo (Thurow 1975). Por último, la sobre-educación puede además verse como una compensación por la falta de otra forma de capital humano, como habilidad, formación en el empleo o experiencia.

Como en otros países europeos, el fenómeno de sobre-educación en España parece haber crecido en importancia desde mediados de los ochenta. En la últimas dos décadas, el mercado laboral español ha experimentado una significativa expansión de la educación post-obligatoria. Esta tendencia, se ha visto especialmente reforzada por la mayor demanda femenina de educación formal, especialmente en lo que se refiere a educación terciaria/universitaria. Como consecuencia de esto, la proporción de población con educación superior está en torno al 75% de la media de la OCDE. Sin embargo, cuando comparamos los logros educativos para las generaciones de 25-34 y 55-64 años, la proporción de gente que ha finalizado estudios terciarios en el primer grupo es 4.5 veces superior al segundo grupo (Dolado et al. 2000). No obstante, a pesar de esta enorme expansión educativa, el mercado laboral español ha experimentado una de las

² Actualmente existe una literatura sustancial, tanto para el caso de América del Norte como Europa, referente al desajuste entre los logros educativos del trabajador y los requisitos educativos del puesto de trabajo. Su importancia queda reflejada en los trabajos de Green, McIntosh y Vignoles (1999), Oosterbeek (2000), Hartog (2000), y (Borghans y de Grip, 2000a).

³ Trabajos recientes, sin embargo, ponen en evidencia estos resultados. Dolton y Vignoles (2000) encuentran que un porcentaje significativo (30%) de trabajadores con educación universitaria están sobre-educados seis años después de completar sus estudios. Utilizando datos para Alemania, Büchel y Mertens (2004) encuentran que los trabajadores sobre-educados tienen una menor probabilidad de experimentar una mejora ocupacional y/o salarial que aquellos trabajadores empleados en un puesto de trabajo que se ajusta a su nivel educativo.

mayores tasas de desempleo del conjunto de la OCDE, siendo las mujeres y los jóvenes los más afectados, con unas tasas de desempleo de hasta 10 y 20 puntos porcentuales superiores a la media. Parece, por tanto, que el aumento en la oferta de mano de obra altamente cualificada no ha ido acompañada de un incremento equivalente en la oferta de trabajos cualificados, dando origen al fenómeno de sobre-educación.

Uno de los trabajos recientes relacionados con el fenómeno de sobre-educación en España, Alba-Ramírez y Blázquez (2004), da evidencia a lo predicho por la teoría de la movilidad ocupacional. Como se pone de manifiesto en dicho trabajo, la sobre-educación en España puede ser tan sólo un fenómeno de carácter transitorio, ya que los trabajadores sobre-educados cuya educación o formación está altamente relacionada con el trabajo que desempeñan, tienen una mayor probabilidad de ser promocionados dentro de la empresa o de cambiar de puesto de trabajo en un período de tres o cuatro años si sus expectativas de promoción no se han satisfecho. Por otro lado, otros trabajos encuentran evidencia del llamado «efecto desplazamiento» de los trabajadores menos educados por los más educados, de manera que la incorporación al mercado laboral en España se ha venido caracterizando en los últimos años por trabajadores jóvenes que aceptan empleos por debajo de sus logros educativos (Dolado et al. 2000). Además este efecto desplazamiento parece utilizarse para explicar las altas tasas de desempleo de los trabajadores menos educados.

3. CAMBIOS LINGÜÍSTICOS EN CATALUÑA

Al igual que otros países europeos, España se caracteriza por una gran diversidad lingüística. En conjunto, alrededor de un 40% de la población española vive en áreas donde hay dos lenguas oficiales. El castellano es la lengua oficial en la totalidad del territorio español, y el catalán, gallego, y vasco comparten co-oficialidad con el castellano en sus respectivos territorios⁴.

⁴ El Catalán es lengua oficial en Cataluña (6'995, 206 habitantes 2005), Comunidad Valenciana (4'692,449 habitantes), y Baleares (983,131). El Gallego es lengua oficial en Galicia (2'762.198) y el Vasco es lengua oficial en el País Vasco (2'124,846) y el norte de Navarra. Estas regiones representan el 39.81% de la población española (44'108,530). Otras lenguas son el Asturiano ó Bable (con alrededor de 600,000 hablantes no son lenguas oficiales dentro de su territorio, Asturias y norte de Castilla y León) y el Aranés (lengua occitana, lengua oficial en el Valle de Arán, con 9,100 habitantes).

Durante la mayor parte de su historia, el catalán ha sido lengua oficial en Cataluña. Sin embargo, desde los años 40 a los 70, durante el régimen de Franco, el castellano fue declarado como la única lengua oficial en Cataluña (y en todo el territorio nacional), mientras que el catalán quedó relegado al ámbito privado. Esta represión, junto con la masiva llegada a Cataluña de inmigrantes castellano-hablantes, explica que una proporción importante de los habitantes de Cataluña no dominen la lengua catalana, incluso aunque sea su lengua nativa.

Después del régimen franquista, el catalán recobró su carácter oficial⁵; y la Generalitat, reestablecida de nuevo, comenzó a incentivar el uso y propagación del catalán tanto en el ámbito público como privado. La política de normalización, promulgada a principios de los ochenta, extendió el uso del catalán a los campos de la educación, la administración pública y los medios de comunicación públicos, de manera que el catalán fue reemplazando de manera progresiva al castellano como principal lengua en el ámbito de la educación primaria y secundaria.

La Figura 1 muestra la evolución del porcentaje de gente que habla catalán en las tres principales regiones catalano-hablantes. En Cataluña, la única región con políticas públicas activas en materia lingüística, se observa una tendencia creciente en el porcentaje de gente que habla catalán desde mediados de los setenta en adelante. En la Comunidad Valenciana y Baleares, donde las políticas públicas para promocionar el uso del catalán fueron menos activas, esta lengua comparte oficialidad con el castellano.

Esta política de normalización en Cataluña se divide en dos períodos principales. El primer período, comprendido de 1980 a 1990, y donde las escuelas jugaron un papel clave en su difusión, se caracterizó por una sensibilización de la población y un aumento en el conocimiento de dicha lengua. En 1980 se creó la Dirección General de Política Lingüística, y en 1983 el Parlamento Catalán aprobó por unanimidad el Proyecto de Ley de Normalización Lingüística. Esta ley constituyó el proceso de consolidación del catalán, especialmente en los campos de educación, la administración pública y los medios públicos de comunicación. Fue este instrumento legal lo que permitió la formulación de un sistema educativo lingüístico caracterizado por un reemplazamiento progresivo del castellano por el catalán como principal lengua de

⁵ Una información más detallada sobre este aspecto puede verse en los artículos en materia lingüística de la Constitución Española y del Estatuto de Autonomía de Cataluña, recogidos en el Apéndice A.1.

enseñanza en la educación primaria y secundaria, dejando al castellano relegado a un segundo plano. En este sentido, tanto los servicios municipales catalanes en materia lingüística, como los nuevos medios de comunicación en catalán creados por la Generalitat han actuado como elementos clave para la difusión del uso del catalán.

La segunda etapa, comprendida desde 1990 hasta la actualidad, se ha centrado en la expansión de carácter sectorial y territorial. La legislación previa fue refordada, de manera que en 1998 se aprobó una nueva ley, el Plan General de Normalización Lingüística. Este plan estratégico, puesto en marcha por la Generalitat, afecta a organizaciones de todos los ámbitos (público y privado) y sectores (incluido el ámbito científico y técnico) con el objeto de promocionar el uso del catalán tanto entre la sociedad como en el funcionamiento de la administración y la instituciones públicas. La principal novedad con respecto a la Ley de 1983 está en la presencia del catalán en dominios que hasta entonces no habían sido regulados, como los medios de comunicación privados, la industria de la cultura y la esfera socio-económica. En el primero de estos ámbitos, el Plan garantiza un determinado porcentaje del tiempo de difusión en catalán; en el segundo, se permite a la Generalitat establecer cuotas sobre la distribución cinematográfica y exposiciones; y en el tercero se regula el uso del catalán en las empresas públicas, información y reclamaciones, información al cliente, y actividades profesionales y publicitarias.

El Plan además impone a la Generalitat la obligación de desarrollar instrumentos para la planificación general del lenguaje a través de programas que permitan establecer los objetivos y las medidas más convenientes en cada momento, así como la evaluación de los resultados. En este sentido, los censos lingüísticos de 1986, 1991 y 1996 muestran avances importantes en las cuatro áreas del conocimiento del lenguaje (entender, hablar, leer y escribir), aunque todavía exista una cierta disparidad en los correspondientes porcentajes⁶. Los sectores en los que la política de normalización ha tenido más éxito son aquellos en los que la Generalitat tenía competencias exclusivas, mientras que ha sido menos significativo en aquellos que no estaban regulados, como el ámbito privado y socio-económico. A diferencia del País Vasco, donde los porcentajes de gente que lee, habla y escribe vasco aumentaron en proporciones similares, en Cataluña el aumento más significativo se observa en el porcentaje de los que escriben catalán (Casesnoves, Turell y Sankoff 2006) Sin embargo, el aumento en el uso del catalán no ha afectado al uso del castellano: los castellano-hablantes han mejorado su conocimiento de la lengua catalana pero manteniendo, al mismo tiem-

⁶ Según el censo de 1996 un 94.97% de los individuos entienden catalán, mientras que tan sólo el 45.84% pueden escribirlo.

po, el uso del castellano (Vila y Vial 2001). Además, como señala Woolard (1989), los individuos no nacidos en Cataluña tienen una actitud más favorable a aprender catalán que los nacidos en Cataluña. A pesar de esto, desde 1983 se ha producido un significativo aumento en el conocimiento del catalán en el conjunto de Cataluña, en todas las esferas y entre todos los grupos de edad. En la próxima sección proporcionamos evidencia descriptiva de las tasas de sobre-educación por grado de conocimiento del catalán y para distintas características demográficas.

4. DATOS

En este trabajo utilizamos dos muestras aleatorias de 250,000 individuos para los censos de los años 1991 y 1996⁷, procedentes de los Institutos Nacional y Catalán de Estadística (INE-IDESCAT). Estas bases de datos contienen información sobre características personales de los individuos tales como el género, edad, estado civil, educación, lugar de residencia, lugar de nacimiento, número de años en Cataluña, ocupación y grado de conocimiento del catalán. Además, combinamos esta información con datos municipales, con el objeto de capturar los posibles efectos de residir en áreas con altas tasas de empleo o con un amplio conocimiento del catalán.

Restringimos el análisis a los miembros principales del hogar, es decir, padres e hijos, con edades comprendidas entre 16 y 60 años, nacidos en España y que formen parte de la población activa, de manera que la muestra final se reduce a 96,863 individuos para 1991, y 96,985 para 1996. Por otro lado, consideramos a un individuo como sobre-educado cuando sus años de educación exceden los años medios de educación de la ocupación más una desviación estándar⁸. Además, a los individuos se les

⁷ Lamentablemente, los datos lingüísticos del censo de 2001 han sido objeto de controversia y no son de acceso general, por lo que el análisis se limita a estos dos años que corresponden al periodo de transición lingüística en Cataluña, alrededor de una década después de la política de Normalización. Estos dos años corresponden también el periodo previo a las grandes migraciones internacionales que recibió Cataluña durante el cambio de siglo. Cabe, finalmente, señalar que a diferencia del censo de 1991, que comprende todo el territorio español, el censo de 1996 sólo se realiza para Cataluña.

⁸ En el Apéndice A.2 se describe con más detalle el proceso de selección muestral, mientras que el Apéndice A.3 contiene la definición de sobre-educación, así como del resto de las variables de análisis.

pregunta sobre su grado de conocimiento de catalán: entender, hablar, leer y escribir. No obstante, con el objeto de controlar por el posible sesgo ocasionado por el hecho de que los individuos revelen un nivel de conocimiento por encima del que realmente poseen, utilizamos una definición rigurosa sobre el nivel de conocimiento del catalán (Charette y Meng, 1994), de manera que consideramos un nivel básico cuando los individuos declaran entender, sólo hablar, o sólo leer catalán; intermedio cuando leen y hablan; y superior cuando además saben escribirlo.

En la Tabla 1 presentamos la estadística descriptiva para todas las variables incluidas en nuestro análisis, para ambos sexos y para los censos de 1991 y 1996 por separado. Como puede observarse, la incidencia de la sobre-educación oscila en ambos casos entre el 9% y 15% del total de la población activa. Los resultados revelan, además, que las diferencias por género en las tasas de sobre-educación no parecen sistemáticas: mientras que en 1991, la sobre-educación es menos probable entre las mujeres, en 1996 se observa la situación contraria. Sin embargo, puede apreciarse como los porcentajes de sobre-educación se reducen con el tiempo tanto para hombres como para mujeres, hecho que puede asociarse al aumento en los logros educativos observados en el conjunto de la población para ese período. En particular, los resultados muestran que el nivel medio de educación para los adecuadamente educados aumenta del 7.7% y 8.6% al 8.5% y 9.3% para hombres y mujeres respectivamente. Por otro lado, puede observarse que los trabajadores sobre-educados (hombres y mujeres) suelen ser más jóvenes que aquellos ocupados en empleos que se ajustan a su nivel educativo. En conclusión, parece que el desajuste educativo afecta principalmente a los jóvenes y tiende a disminuir en el tiempo como consecuencia de una mejora en los niveles educativos de los trabajadores.

Los resultados de la estadística descriptiva confirman, por otro lado, la existencia de una relación positiva entre la incidencia del fenómeno de sobre-educación y el conocimiento del catalán. Por un lado, es más probable encontrar individuos con un conocimiento intermedio y superior del catalán entre los sobre-educados que entre los individuos adecuadamente educados. Y por otro lado, las tasas de sobre-educación son mayores entre los individuos, hombres y mujeres, que saben catalán. Se observa, además, que el conocimiento medio de catalán aumenta con el tiempo y que las mujeres suelen ser más expertas que los hombres en el conocimiento de esta lengua.

Los resultados ponen además de manifiesto, por un lado, que el riesgo de sobre-educación es mayor entre los individuos empleados en el sector servicios, y por otro que las mujeres tienen una mayor probabilidad de ocupar un trabajo en este sector, lo que puede explicar, en parte, las diferencias de género en la incidencia de la sobre-educación.

Encontramos, además, que hay más hombres casados que mujeres, y que el porcentaje de población afectado directamente por el proceso de normalización es mayor para las mujeres que para los hombres y que, para ambos sexos, se observa una tendencia creciente en el tiempo: para los hombres el porcentaje de afectados por el proceso de normalización ha aumentado del 3.3% en 1991 al 10.2% en 1996, mientras que para las mujeres este porcentaje varía del 3.0% al 12.7%.

Los datos revelan, por otra parte, que la mayor parte (80%) de la población en Cataluña se concentra en la provincia de Barcelona, aunque con el tiempo parece que esta proporción tiende a disminuir, y que Girona, Tarragona y Lleida siguen, por este orden, a Barcelona como principales regiones de residencia. Por otro lado, los datos a nivel de municipio confirman la tendencia creciente del uso del catalán, así como una mejora en el grado de conocimiento de esta lengua. Finalmente, se observa una tendencia creciente de la presencia del sector servicios en la economía catalana, así como un descenso en las tasas de empleo para ambos sexos.

Aproximadamente un tercio de la muestra consiste en gente nacida fuera de Cataluña, de los cuales más de dos tercios proceden de Andalucía, mientras que una proporción muy baja procede de Valencia, Baleares o La Franja (las regiones de Aragón colindantes con Cataluña) donde tanto catalán como castellano son lenguas oficiales. Los datos revelan además una reducción de la proporción de individuos que no han nacido en Cataluña, así como una disminución del porcentaje, tanto de hombres como de mujeres, nacidos en Andalucía. Por otro lado, la mayoría de los individuos procedentes de otras regiones de España llegaron a finales de los sesenta y tienen una media de edad de entre 23 y 25 años en 1991, y de entre 27 y 28 años en 1996, y aproximadamente un tercio de ellos tenía 10 años o menos cuando llegaron a Cataluña.

En resumen, el análisis descriptivo parece sugerir que el saber catalán aumenta la probabilidad de sobre-educación. A partir de estos primeros resultados cabría pensar, como suele argumentarse en numerosos debates públicos sobre la política lingüística en Cataluña, que aquellos individuos que no son expertos en el conocimiento del catalán puedan sentirse desanimados y tengan menos incentivos a educarse. En la siguiente sección veremos, sin embargo, como este planteamiento cambia cuando tenemos en cuenta el efecto de otros factores y controlamos por el carácter endógeno del conocimiento del catalán.

5. ESTIMACIÓN

En esta sección llevamos a cabo un análisis en profundidad del problema de sobre-educación en Cataluña, así como de su relación con el conocimiento del catalán. Para este propósito estimamos, en primer lugar, un modelo probit estándar para los años 1991 y 1996, y para hombres y mujeres por separado, donde el saber catalán se considera como un factor explicativo exógeno a la hora de estimar la probabilidad de sobre-educación. En la estimación de estos modelos probit incluimos como variables explicativas las siguientes características personales: años de educación y su cuadrado, edad y su cuadrado, interacción entre edad y educación, estado civil, ocupación y sector de actividad. Además, controlamos por la situación del mercado laboral local incluyendo como variables explicativas la tasa de empleo del municipio, la proporción de empleados en el sector servicios en el municipio, y un conjunto de variables dicotómicas para controlar por la región de residencia. Además, cabe señalar que realizamos estimaciones alternativas para distintos niveles de conocimiento del catalán: intermedio (leer y hablar) y superior (escribir)⁹.

Los resultados obtenidos de la estimación del modelo probit estándar serían correctos siempre y cuando el saber catalán fuese un factor explicativo exógeno, por ejemplo si se tratase de un atributo de carácter étnico (como se ha asumido en trabajos recientes sobre la economía del lenguaje (Becker 1957 y Reynauld y Marion 1972; véase Grin 2003)). Sin embargo, dichos resultados estaría sesgados si hubiese determinados atributos o factores que afectasen al mismo tiempo al hecho de saber catalán y a la probabilidad de estar sobre-educado. Incluso en esta estimación de variable dependiente limitada el efecto del conocimiento del catalán en la sobre-educación puede en realidad estar capturando el efecto de las variables que determinan el conocimiento de catalán. Para controlar por este posible sesgo de selección muestral, procedemos a estimar la probabilidad de que un individuo esté sobre-educado condicionada a saber catalán¹⁰.

Para este propósito seguimos el enfoque de Willis y Rosen (1979) y procedemos con una estimación en dos etapas (modelo probit bivariado). En la primera etapa esti-

⁹ Los resultados de estas estimaciones aparecen recogidos en la Tabla 2.

¹⁰ De nuevo realizamos estimaciones separadas para hombres y mujeres, y para los Censos de 1991 y 1996.

mamos la probabilidad de saber catalán incluyendo como factores explicativos, a parte de los incluidos en la ecuación de sobre-educación, los siguientes: porcentaje de individuos nacidos en Cataluña, porcentaje de individuos que escriben catalán en el municipio, una variable dicotómica que indica si el individuo se ha visto afectado por el proceso de normalización, una variable dicotómica que indica si el individuo tenía menos de 10 años cuando llegó a Cataluña, años desde que emigró, una interacción entre estas dos últimas variables, y un conjunto de variables dicotómicas que identifican la región de origen: Andalucía, Valencia, Baleares o La Franja¹¹. Todas estas variables, que tratan de recoger los principales factores que determinan el grado de conocimiento del catalán, son incluidas en la ecuación de selección pero no en la ecuación de sobre-educación actuando, por tanto, como instrumentos válidos para la indentificación del modelo probit bivariado (Maddala, 1983).

Dado que el hecho de nacer fuera de la región actual de residencia determina que un individuo sea considerado como «inmigrante» y, por tanto, distinto a un «nativo» en diferentes aspectos más allá del conocimiento de la lengua local, resulta bastante razonable utilizar variables de origen como instrumentos válidos en la estimación de nuestro modelo en dos etapas. No obstante, hemos de tener en cuenta que en nuestro análisis estamos comparando ciudadanos españoles entre los que no existen diferencias en términos de raza, religión o situación legal. Además, no puede decirse que los individuos que no han nacido en Cataluña puedan considerarse exactamente como «recién llegados», ya que muchos de ellos llegaron cuando tenían una media de entre 23 y 25 años, y un tercio de ellos cuando tenían menos de 10 años, de manera que, independientemente de su grado de conocimiento de catalán, estos individuos son considerados como catalanes. Por otro lado, una característica adicional, y que distingue el caso de Cataluña con respecto a otras economías, es que no todos los nativos saben la lengua local (catalán), pero todos los individuos saben la lengua del lugar de procedencia de los inmigrantes (castellano). Todas estas particularidades hacen que las variables que hemos seleccionado como instrumentos sólo estén correlacionados con el conocimiento del catalán pero no con la probabilidad de sobre-educación.

Una vez estimados los parámetros de la primera etapa (ecuación de selección), procedemos con la estimación de la segunda etapa, la probabilidad de estar sobre-

¹¹ Los resultados de la estimación de la primera etapa son muy parecidos a los obtenidos en Rendon (2006), y están disponibles bajo previa petición del lector.

educado condicionada a un determinado grado de conocimiento del catalán. Los resultados de las estimaciones de esta segunda etapa aparecen recogidos en la Tabla 3. Como puede observarse en esta tabla, existe un efecto positivo del nivel educativo sobre la probabilidad de estar sobre-educado, lo cual no resulta sorprendente ya que en los últimos años el mercado laboral español se ha venido caracterizando por una notable incidencia del fenómeno de sobre-educación entre aquellos individuos con un nivel de educación terciaria/universitaria. Por el contrario, el coeficiente estimado sobre la variable «edad» resulta negativo y significativo, lo que confirma la tendencia observada recientemente en el mercado laboral español de trabajadores jóvenes aceptando empleos para los que están sobre-educados, mientras continúan buscando un empleo más adecuado a su nivel de cualificación.

En relación a las variables explicativas referentes a la región de residencia, nuestros resultados revelan, para el censo de 1991, un efecto negativo y significativo de la tasa de empleo en el municipio sobre la probabilidad individual de estar sobre-educado. Este resultado puede explicarse de la forma siguiente: si el porcentaje de empleados en el municipio está positivamente relacionado con la probabilidad de encontrar un empleo, resulta lógico que cuanto mayor sea el porcentaje de empleados en el municipio, mayor será la probabilidad de que un individuo encuentre un empleo que se ajuste a su nivel de cualificación. Para 1996 encontramos, además, un efecto negativo y significativo de la variable «porcentaje de empleados en el sector servicios en el municipio» sobre la probabilidad de sobre-educación para el caso de los hombres, mientras que para las mujeres este efecto resulta positivo y significativo, lo cual puede explicarse, tal como queda reflejado en los resultados de la Tabla 1, porque por un lado el fenómeno de sobre-educación es más frecuente en el sector servicios, y por otro, porque la proporción de empleo femenino es mayor en el sector servicios que en el conjunto de la economía.

Nótese, además, que el coeficiente de correlación es negativo y significativamente distinto de cero, con valores comprendidos entre -5% y -10%, lo que indica que la ecuación de sobre-educación no es independiente de la ecuación de selección, y que existen factores inobservables que aumentan el grado de conocimiento de catalán y que están negativamente correlacionados con inobservables que aumentan el riesgo de sobre-educación. El test de ratio de verosimilitud confirma que la hipótesis de independencia entre las dos ecuaciones debe rechazarse, lo que sugiere que es imprescindible controlar por la selección muestral a la hora de estimar la probabilidad de sobre-educación.

En la Tabla 4 presentamos las probabilidades predichas de estar sobre-educado en función del grado de conocimiento de catalán. Como señalaban los datos de la esta-

dística descriptiva, para todos los grupos el porcentaje de sobre-educados es considerablemente mayor entre los que saben catalán que entre los que no lo saben. No obstante, este resultado no es muy relevante ya que se obtiene a partir de grupos de individuos con diferentes características. Una comparación válida debería hacerse calculando, para un mismo subgrupo de la población, los efectos discretos individuales y medios de saber catalán, a partir de los parámetros estimados del modelo en dos etapas. Encontramos que el efecto de saber catalán sobre la probabilidad de sobre-educación resulta negativo y significativo a nivel individual para la mayoría de los grupos de análisis, si bien es cierto que la magnitud, en términos absolutos, de estos efectos se reduce cuando incorporamos variables dicotómicas en el análisis de ocupación y sector de actividad. Por otro lado, encontramos también que los efectos discretos medios resultan negativos, lo que sugiere que saber catalán reduce la probabilidad de sobre-educación¹². Así, si comparamos las tasas de sobre-educación entre los que saben catalán (17.62%) y los que no saben (8.93%), podemos pensar que saber catalán aumenta el riesgo de sobre-educación. Por el contrario, los resultados obtenidos del modelo probit bivariado revelan que saber catalán reduce la probabilidad de sobre-educación en 1.85 puntos porcentuales para los individuos que saben catalán, y en 0.84 para aquellos que no lo saben. Estos resultados podrían sugerir la existencia de algún tipo de «efecto sustitución» entre cualificaciones lingüísticas y educación, en la medida en que los trabajadores que saben catalán tienen una mayor probabilidad de ocupar empleos que se ajustan mejor a sus logros educativos. Si bien es cierto, que este mejor ajuste podría ser debido al hecho de que los individuos que saben catalán pueden acceder a más puestos de trabajo, tanto aquellos en los que es necesario saber catalán como aquellos en los que no. En consecuencia, al haber más puestos disponibles para estos individuos, la posibilidad de conseguir un trabajo que se ajuste mejor a sus logros educativos sería mayor. Sin embargo, encontramos que los efectos discretos medios tienen signo negativo (y con mayor valor en términos absolutos que los efectos individuales), pero son claramente no significativos. Esto se debe a que las desviaciones estándar son muy altas, con lo que no podemos rechazar la hipótesis de que estos efectos discretos medios sean significativamente distintos de cero. Esta diferencia, en términos de significatividad, entre los efectos medios y los de un individuo representativo puede deberse a la curvatura de la función de distribución acumulada (desigualdad de Jensen): las desviaciones estándar de los efectos

¹² Las expresiones detalladas de estos efectos discretos medios quedan recogidas en el Apéndice A.4 del trabajo de Blázquez y Rendon (2006).

medios son significativamente mayores a las correspondientes a los efectos individuales, lo que puede explicarse por la existencia de una gran variación de estos últimos entre el conjunto total de individuos de la muestra.

Nótese, además, que el efecto del conocimiento del catalán sobre la probabilidad de sobre-educación es mayor para los hombres que para las mujeres, disminuye con el grado de conocimiento de catalán y con el tiempo, y es mayor para individuos que saben catalán que para los que no lo saben.

Por último, procedemos a contrastar la robustez de los resultados probando con diferentes especificaciones del modelo. Los resultados aparecen reflejados en la Tabla 5, donde se recogen los efectos discretos medios de los modelos probit estándar y probit bivariado, con y sin variables de ocupación y sector de actividad. Para las distintas alternativas, los resultados no varían de manera significativa. Sin embargo, controlar por la endogeneidad de la variable «conocimiento de catalán» tiende a aumentar el valor absoluto de los efectos discretos medios sobre la probabilidad de sobre-educación. Además, para el caso de un nivel superior de conocimiento de catalán, los efectos discretos medios obtenidos de las estimaciones del modelo probit bivariado son, por lo general, menores a los del modelo probit estándar.

En resumen, podemos concluir que el problema de selección muestral es importante, y que no controlar por este posible sesgo de selección conduce a infra-estimar el efecto del catalán sobre la probabilidad de sobre-educación, especialmente cuando nos centramos en un nivel superior de conocimiento. Además, en la mayoría de los casos encontramos que el saber catalán tiende a reducir el riesgo de sobre-educación, aunque dicho efecto no resulta estadísticamente significativo a nivel agregado.

6. CONCLUSIONES

En este trabajo examinamos si existe algún tipo de relación entre cualificaciones lingüísticas y el fenómeno de sobre-educación. El análisis se centra en el mercado laboral catalán, caracterizado por la presencia de dos lenguas oficiales: catalán y castellano. El catalán, inicialmente relegado al uso popular y familiar, se convirtió en lengua co-oficial, junto con el castellano, a principios de los ochenta. Este cambio condujo a un aumento considerable en el uso y conocimiento del catalán, aunque el castellano continuase conservando su importancia en muchas esferas de la comunicación.

Nuestro principal objetivo se centra en contrastar si aquellos individuos con una carencia de cualificaciones lingüísticas pueden tratar de compensar esta deficiencia aceptando empleos para los que están sobre-educados. A primera vista, los resultados

de la estadística descriptiva sugieren que los individuos con un mayor nivel de conocimiento del catalán presentan una mayor probabilidad de estar sobre-educados. Estos primeros resultados parecen sugerir la presencia de un «efecto desincentivador»: los individuos que no dominan la lengua de enseñanza en la escuela sufren un desaliento en proseguir su educación formal, acaban por fracasar en la escuela y, por lo tanto, se educan menos.

Sin embargo, una vez que procedemos a realizar un análisis más detallado y estimamos un modelo que nos permite controlar por diversos factores, encontramos que saber catalán ejerce un efecto negativo sobre la probabilidad de que un individuo esté sobre-educado, lo que parece ir en línea con la hipótesis de «sustitución de cualificaciones»: los individuos que presentan una carencia de cualificaciones lingüísticas, pueden tratar de compensar esta deficiencia adquiriendo mayores niveles de educación formal. Nuestros resultados ponen de manifiesto que, aunque este efecto resulta significativo a nivel individual, y es robusto al carácter endógeno del conocimiento del catalán, el alto grado de heterogeneidad entre los individuos hace que los efectos discretos medios no sean significativos. Complementariamente, nuestros resultados puede reflejar que los individuos que saben catalán pueden acceder a más puestos de trabajo, y por lo tanto, su probabilidad de desajuste educativo es menor que los que no lo saben. Bajo este mecanismo, el conocimiento de la lengua catalana es también una cualificación apreciada en el mercado de trabajo.

En resumen, los resultados obtenidos en este trabajo permite derivar implicaciones importantes relacionadas con políticas en materia lingüística, especialmente en aquellos países que han experimentado cambios en su lengua de enseñanza en la escuela, como el cambio de francés a árabe en Marruecos, de inglés a galés en Gales, de ruso a estonio en Estonia, o de inglés a español en Puerto Rico. En todos estos países, aquellos individuos que no tienen un dominio completo de la lengua de enseñanza pueden, o bien presentar unos menores niveles de educación si la carencia de cualificaciones lingüísticas desincentiva a estos individuos a invertir en educación, o un mayor nivel de educación si tratan de compensar esta deficiencia con otras formas de capital humano. Los dos efectos son, en principio, posibles y pueden ser objeto de debate en materia de política lingüística. Para el caso de Cataluña, nuestros resultados sugieren que el efecto sustitución entre cualificaciones lingüísticas y educación formal predomina ligeramente sobre el efecto desincentivador. Si bien es importante señalar, que el hecho de que el saber catalán reduzca la probabilidad de sobre-educación puede ser simplemente debida a que los individuos que saben catalán pueden acceder a más puestos de trabajo, tanto aquellos en los que es necesario saber catalán como aquellos en los que no.

APÉNDICE

A.1. Estado legal del catalán

Lo recogido a continuación se ha tomado del trabajo de Webber y Strubelli (1991), que a su vez está extraído de artículos de la Constitución Española y del Estatuto de Autonomía de Cataluña relacionados con el uso de las dos lenguas oficiales en Cataluña: castellano y catalán.

Constitución Española (1978). Artículo 3:

1. El castellano es la lengua oficial del Estado. Todos los españoles tienen la obligación de saberlo y el derecho a usarlo.
2. El resto de lenguas españolas serán también oficiales en sus respectivas Comunidades Autónomas, de acuerdo con sus correspondientes Estatutos.
3. La riqueza de las diferentes modalidades lingüísticas en España es un patrimonio cultural que deberá ser objeto de especial respeto y protección.

Estatuto de Autonomía de Cataluña (1979). Artículo 3:

1. El idioma propio de Cataluña es el catalán.
2. El catalán es la lengua oficial en Cataluña, como lo es el castellano sobre todo el Estado español.
3. La Generalitat garantizará el uso normal y oficial de ambas lenguas, tomará las medidas necesarias para asegurar el conocimiento adecuado de ambas, y creará las condiciones necesarias que permitirán una completa igualdad entre ambas lenguas, en relación a los derechos y obligaciones de todos los ciudadanos de Cataluña.

A.2. Selección muestral

La siguiente tabla recoge la importancia de los criterios de selección a la hora de construir la muestra.

	1991	1996
Muestra total	250.000	250.000
Sólo miembros principales del hogar: padres e hijos	-17.654	-17.903
Sólo individuos entre 16 y 60 años	-82.297	-81.770
Sólo ciudadanos españoles	-5.740	-4.745
Sólo si se tiene la variable llegada a Cataluña		-3.788
Sólo individuos pertenecientes a la población activa	-47.421	-44.809
Sólo si se tiene la variable conocimiento de catalán	-25	
Muestra seleccionada	96.863	96.985

A.3. *Definición de las variables*

A continuación explicamos cómo se han construido las variables que vamos a utilizar en el análisis.

Sobre-educación.- Un trabajador se define como sobre-educado si sus años de educación están por encima de la media de la ocupación más una desviación estándar. Se considerará como adecuadamente educado si sus años de estudios están por encima de la media de la ocupación menos una desviación estándar y por debajo de la media más una desviación estándar. Por último, consideraremos como infra-educados a aquellos individuos con un nivel de estudios inferior a la media de la ocupación menos una desviación estándar¹³.

Educación.- El censo proporciona información sobre el máximo nivel de estudios del individuo. Para cada nivel asignamos los años correspondientes de educación.

Edad.- Año del censo, 1991 o 1996, respectivamente, menos el año de nacimiento.

¹³ En la literatura podemos encontrar dos tipos de definiciones de sobre-educación: de tipo objetivo y de tipo subjetivo. Las definiciones de tipo subjetivo se basan en la información proporcionada por el trabajador sobre su nivel de cualificación. Las definiciones de carácter objetivo pueden clasificarse en dos tipos. En el primer caso la sobre-educación se determina comparando los años de educación del trabajador con el nivel educativo medio de la ocupación. El segundo tipo se basa en la comparación entre el nivel educativo del trabajador y los requisitos del puesto de trabajo. En el presente trabajo utilizamos el primer tipo de definiciones de carácter objetivo.

Normalización.- Variable dicotómica que toma valor 1 si el individuo tenía menos de 12 años en 1984 y 0 en caso contrario.

Casado.- Variable dicotómica que toma valor 1 si el individuo está casado en el momento del censo y 0 en caso contrario.

Variabes de residencia.- El censo proporciona información sobre el municipio y provincia de residencia del individuo. Con esta información podemos construir variables dicotómicas para identificar las provincias de Lleida, Girona y Tarragona.

Variabes de origen.-El censo proporciona información sobre el municipio y provincia de nacimiento del individuo. Con esta información podemos construir variables dicotómicas para identificar a los individuos nacidos en Cataluña, Andalucía, Valencia, Baleares y La Franja.

Años desde la Migración.- El censo proporciona información sobre el año de llegada a Cataluña. A partir de esta información podemos saber el número de años desde que llego a Cataluña (años del censo menos el año de llegada). Además construimos una variable dicotómica que identifica a aquellos individuos que llegaron a Cataluña con menos de 10 años.

Variabes del municipio.- Utilizamos las variables de residencia para asignar a cada individuo la información correspondiente al municipio.

Ocupación.- (1) Agricultura (grupo de referencia), (2) Industria, (3) Comercio, Servicios y Profesionales.

Actividad.- (1) Agricultura (grupo de referencia), (2) Industria, (3) Comercio, (4) Servicios.

REFERENCIAS BIBLIOGRÁFICAS

ACEMOGLU, D. (2002), «Technical Change, Inequality, and the Labour Market», *Journal of Economic Literature* 40, 7-72.

ALBA-RAMÍREZ, A. (1993), «Mismatch in the Spanish Labour Market. Overeducation?», *Journal of Human Resources* 28, 259-278.

ALBA-RAMÍREZ, A., y BLÁZQUEZ, M. (2004). «Types of Job Match, Overeducation and Labour Mobility in Spain» en F. Büchel, A. de Grip y A. Mertens (eds), *Overeducation in Europe: Current Issues in Theory and Policy*, Edward Elgar, Cheltenham, UK.

BARRET, A., BERGIN, A. y DUFFY, D. (2005), «The Labour Market Characteristics and Labour Market Impacts of Immigrants in Ireland», IZA DP 1553.

- BATTU, W. y SLOANE, P.J. (2002), «Overeducation and Ethnic Minorities in Britain», IZA DP 650.
- BÜCHEL, F., y MERTENS, A. (2004), «Overeducation, Undereducation and the Theory of Career Mobility», *Applied Economics* 36.
- BECKER, G. (1957), *The Economics of Discrimination*, Chicago University Press, Chicago.
- BLÁZQUEZ, M., y RENDON, S. (2006), «Overeducation in Multilingual Economies: Evidence from Catalonia», ITAM Working Paper, N° 06-07.
- BORGHANS, L., y DE GRIP, A. (2000), *The Overeducated Worker? The Economics of Skill Utilization*, Borghans, L., y de Grip, A. (eds).
- BORJAS, G. (1999), «The Economic Analysis of Immigration» en O. Ashenfelter y D. Card (eds), *Handbook of Labor Economics* vol 3A, chapter 28, North Holland, Amsterdam, pp. 1697-1760.
- BRATSBERG, B. y RAGAN, J.F. (2002), «The Impact of Host-country Schooling on Earnings. A Study of Male Immigrants in the United States». *Journal of Human Resources* 37.
- CASESNOVES FERRER, R., TURELL, M.T. y SANKOFF, D. (2004), «La base démolinguistique pour évaluer l'aménagement linguistique dans un contexte bilingue». X Congrés Linguapax. Diversitat lingüística, sostenibilitat i pau. Fòrum Diàlegs, Barcelona.
- CHARRETTE, M. y MENG, R. (1994), «Explaining Language Proficiency. Objective versus self-assessed measures of literacy», *Economic Letters* 44, 313-321.
- CHISWICK, B. y MILLER, P. (2002), «Immigrant Earnings: Language Skills, Linguistic Concentrations and the Business Cycle», *Journal of Political Economy* 15, 31-57.
- DEKKER, R. d.A y HEIJKE, H. (2002), «The Effects of Training and Overeducation on Career Mobility in a Sequential Labour Market», *International Journal of Manpower* 23.
- DOLADO, J., FELGUEROSO, F. y JIMNEO, J. (2000), «Youth Labour Market in Spain: Education, Training and Crowding-out», *European Economic Review* 44.
- DOLTON, P. y VIGNOLES, A. (2000), «The incidents and Effects of Overeducation in the UK Graduate Labour Market», *Economics of Education Review* 19, 179-198.
- DUPUY, A. y DE GRIP, A. (2002), «Do Large Firms Have More Opportunities to Substitute between Skill Categories than Small Firms?», Aarhus Centre for Labour Market and Social Research, Working Paper 02-01.
- FREEMAN, R. (1976), *The Overeducated American*, Academic Press, New York.
- GARCÍA-SERRANO, C. y MALO, M. (1996), «Desajuste Educativo y Movilidad Laboral en España», *Revista de Economía Aplicada* (11), 105-131.
- GREEN, C., KLER, P. y LEEVES, G. (2004), «Overeducation and Assimilation of Recently Arrived Immigrants: Evidence from Australia», University of Queensland, WP 4.

- GRIN, F. (2003), «Language planning and economics», *Current Issues in Language Plannig* 4(1), 1-66.
- GROOT, W. (1993), «Overeducation and the Returns to Enterprise-related Training», *European Economic Review* 12, 299-309.
- (1996), «The Incidence of, and Returns to Overeducation in the UK», *Applied Economics* 28, 1345-1350.
- GROOT, W., y MAASSEN VAN DEN BRINK, H. (1997), «Allocation and the Returns to Overeducation in the United Kingdom», *Education Economics* 5, 169-183.
- HARTOG, J. (2000), «Overeducation and Earnings: Where Are We, Where Should We Go?», *Economics of Education Review* 19, 131-147.
- JOVANOVIC, B. (1979), «Job Matching and the Theory of Turnover», *Journal of Political Economy* 87, 972-990.
- LINSLEY, I. (2005), «Overeducation in the Australian Labour Market: Its Incidence and Effects». The University of Melbourne, WP 939.
- MADDALA, G.S. (1983), *Limited-dependent and qualitative variables in econometrics*, Cambridge University Press, Cambridge.
- NIESING, W., VAN PRAAG, B. Y VEENMAN, J. (1994), «The Unemployment of Ethnic Minority Groups in the Netherlands», *Journal of Econometrics* 61.
- OOSTERBEEK, H. (2000), «Introduction to Special Issue on Over-schooling», *Economics of Education Review* 19.
- RAYNAULD, A. y MARION, P. (1972), «Une analyse économique de la disparité inter-ethnique des revenus», *Revue Économique* 23, 1-19.
- RENDON, S. (2006), «The Catalan Premium: Work and Language in Catalonia», en prensa en *Journal of Population Economics*.
- ROBST, J. (1995), «Career Mobility, Job Match, and Overeducation», *Eastern Economic Journal*.
- ROSEN, S. (1972), «Learning and Experience in the Labour Market», *Journal of Human Resources* 7, 326-342.
- SCHAAFSMA, J. y SWEETMAN, A. (2001), «Immigrant Earnings: Age at Immigration Matters», *Canadian Journal of Economics* 34.
- SICHERMAN, N. (1991), «Overeducation in the Labor Market», *Journal of Labor Economics* 9, 101-122.
- SICHERMAN, N., y GALOR, O. (1990), «A Theory of Career Mobility», *Journal of Political Economy* 98, 160-192.
- SPENCE, M. (1973), «Job Market Signalling», *Quarterly Journal of Economics* 87, 355-374.
- THUROW, L. (1975), *Generating Inequality: Mechanisms of Distribution in the U.S Economy*, New York: Basic Books, New York.
- VILA, F.X. y VIAL, S. (2001), Escola i 'us. les pràctiques lingüístiques de l'alumnat de 2n nivell de cicle superior d'educació primària de Catalunya en situacions quasi-spontànies, Mimeo

- WEBBER, J. y Strubell i Trueta, M. (1991), *The Catalan Language. Progress towards Normalisation*, The Anglo-Catalan Society, Exeter.
- WIELING, M. y BORGHANS, L. (2001), «Discrepancies between supply and demand and adjustment processes in the labour market», *Labour* 15, 33-56.
- WOOLARD, K.A. (1989), *Double Talk. Bilingualism and the politics of ethnicity in Catalonia*, Stanford University Press cop., Stanford, California.

Tabla 1. Estadística descriptiva

Censo Sexo	1991		1996	
	Homb.	Muj.	Homb.	Muj.
Educación				
%Adecuadamente educados	71.5	75.9	81.2	80.7
%Sobre-educados	14.4	11.6	9.7	10.9
Nivel de estudios de los:				
Adecuadamente educados	7.7	8.6	8.5	9.3
Sobre-educados	12.0	12.8	12.8	13.6
Edad de los:				
Adecuadamente educados	37.8	34.4	38.4	36.2
Sobre-educados	34.3	31.5	36.3	33.5
%Gente que lee y habla catalán entre los:				
Adecuadamente educados	64.9	75.2	75.0	83.8
Sobre-educados	77.1	82.1	88.9	92.8
%Gente que escribe catalán entre los:				
Adecuadamente educados	37.1	51.5	46.0	60.7
Sobre-educados	54.0	65.2	67.6	80.6
%Sobre-educados entre los que:				
Leen y hablan catalán	17.6	13.0	11.5	12.1
No leen ni hablan catalán	8.9	7.7	4.2	4.6
Escriben catalán	21.2	15.1	14.0	14.3
No escriben catalán	10.5	8.1	5.9	5.4
%Sobre-educados en el total de la muestra	14.4	11.6	9.7	10.9
%Sobre-educados en el sector servicios	17.7	12.8	11.6	12.0
%Mujeres		34.5		37.3
%Mujeres trabajando en el sector servicios		56.6		57.0
%Casados	69.6	59.5	67.0	59.9
%Afectados por la política de Normalización	3.1	3.9	10.2	12.7
Residencia				
%Que viven en Barcelona	86.3	77.7	75.0	75.8
%Que viven en Girona	8.7	8.9	9.1	9.7
%Que viven en Lleida	5.9	5.5	6.3	5.7
%Que viven en Tarragona	9.1	7.9	9.6	8.8
Municipio				
%Que escriben catalán en el municipio	39.8	40.5	46.0	46.4
%Nacidos en Cataluña en el municipio	67.3	67.7	68.5	68.6
%Que trabajan en el sect. servicios en el municipio	51.6	52.6	57.2	58.2
%Empleados sobre población del municipio	37.2	37.5	36.4	36.4
Origen				
%Nonacidos en Cataluña	37.6	30.2	31.0	24.9
%Nacidos en Andalucía	24.1	18.3	19.5	14.4
%Nacidos en Valencia-Baleares	1.2	1.2	1.1	1.1
%Nacidos en La Franja	0.5	0.5	0.4	0.4
%Que llegaron con menos de 10 años	11.0	11.1	10.0	9.6
Años desde que migraron	25.0	23.6	28.4	26.8

Tabla 2. Efectos discretos del conocimiento del catalán sobre la probabilidad de sobre-educación. Errores estándar en letra pequeña

Conocimiento de catalán	Leer y hablar				Escribir			
	1991		1996		1991		1996	
	Homb.	Muj.	Homb.	Muj.	Homb.	Muj.	Homb.	Muj.
Censo								
Sexo								
Básico	-2.14	-4.88	-0.15	-2.06	-1.78	-3.63	-0.09	-1.06
Indiv. representativo	0.21	0.45	0.03	0.35	0.15	0.29	0.02	0.16
Media	-3.54	-4.44	-2.75	-5.67	-3.17	-3.52	-1.89	-3.48
	3.51	5.20	3.52	6.52	3.18	4.31	2.45	4.11
Variables de ocupación y de actividad								
Indiv. representativo	-0.30	-0.65	0.00	-0.29	-0.25	-0.52	0.00	-0.15
	0.07	0.15	0.00	0.09	0.05	0.10	0.00	0.04
Media	-1.33	-2.02	-1.04	-3.16	-1.17	-1.73	-0.72	-1.88
	1.61	2.72	1.46	3.69	1.40	2.35	1.01	2.22

Tabla 3. Ecuación de sobre-educación por grado de conocimiento del catalán.
Errores estándar en letra pequeña

Conocimiento de catalán	Leer y hablar				Escribir			
	1991		1996		1991		1996	
	Homb.	Muj.	Homb.	Muj.	Homb.	Muj.	Homb.	Muj.
Censo								
Sexo								
Educación	1.13	0.97	3.02	1.92	1.13	0.97	3.02	1.92
	0.02	0.03	0.06	0.07	0.02	0.03	0.06	0.07
Educación ²	-2.82	-2.09	-10.54	-6.20	-2.82	-2.11	-10.54	-6.20
×10 ⁻²	0.06	0.09	0.24	0.27	0.06	0.09	0.24	0.27
Edad	-0.18	-0.17	-0.36	-0.45	-0.19	-0.18	-0.37	-0.46
×10 ⁻¹	0.07	0.11	0.09	0.12	0.07	0.11	0.09	0.12
Edad ²	0.26	0.43	0.48	0.54	0.26	0.45	0.48	0.55
×10 ⁻³	0.09	0.14	0.12	0.15	0.09	0.14	0.12	0.15
Edad × Educación	-10.01	-23.10	-6.66	-7.36	-9.77	-22.81	-6.46	-6.93
×10 ⁻⁴	3.07	4.21	4.82	6.21	3.07	4.21	4.82	6.21
Casado	0.70	-0.00	-0.52	-0.15	0.70	0.01	-0.51	-0.14
×10 ⁻¹	0.24	0.30	0.27	0.28	0.24	0.30	0.27	0.28
Residente en Lleida	0.73	0.27	-0.52	-0.03	0.78	0.31	-0.51	-0.01
×10 ⁻¹	0.39	0.54	0.48	0.56	0.39	0.54	0.48	0.56
Residente en Girona	0.44	-0.28	0.05	-0.48	0.45	-0.24	0.06	-0.46
×10 ⁻¹	0.36	0.51	0.45	0.51	0.36	0.51	0.45	0.51
Residente en Tarragona	-0.04	-1.21	-1.54	-0.44	-0.03	-1.20	-1.52	-0.41
×10 ⁻¹	0.34	0.52	0.42	0.49	0.34	0.52	0.42	0.49
%Empleados ^a	-1.69	-1.47	-0.64	-0.07	-1.68	-1.50	-0.63	-0.07
	0.47	0.68	0.47	0.56	0.47	0.68	0.47	0.56
%Empleados en servicios ^a	-0.03	0.16	-0.24	0.37	-0.03	0.15	-0.25	0.36
	0.08	0.12	0.10	0.12	0.08	0.12	0.10	0.12
Variables dicotómicas								
Sector actividad	✓	✓	✓	✓	✓	✓	✓	✓
Ocupación	✓	✓	✓	✓	✓	✓	✓	✓
ρ	-0.06	-0.11	-0.05	-0.10	-0.05	-0.10	-0.04	-0.08
	0.01	0.02	0.02	0.03	0.01	0.02	0.02	0.02
LRT($\rho = 0$)	16.42	9.21	6.21	12.72	13.42	26.43	5.38	13.40

^a Porcentaje en el municipio.

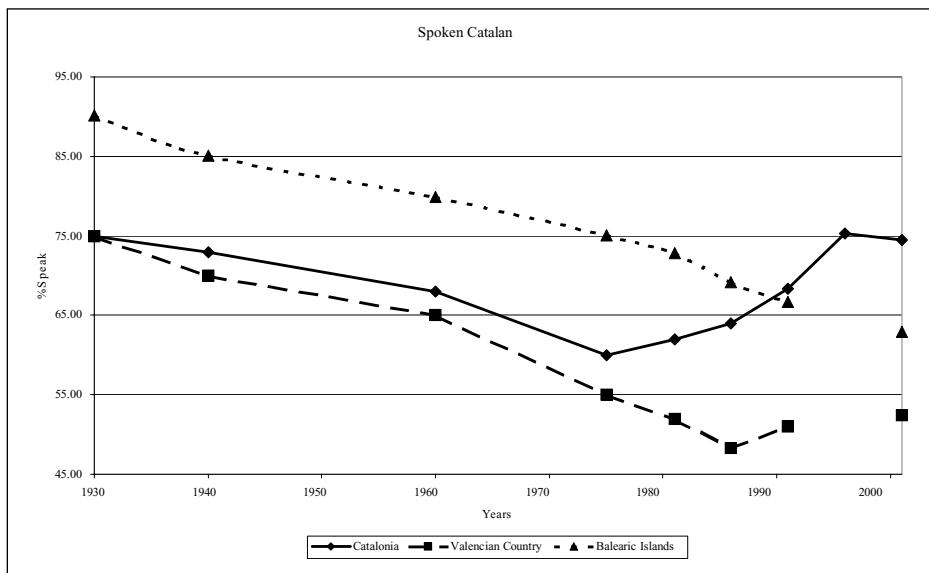
Tabla 4. Probabilidad de sobre-educación. Efectos discretos medios (en%) por grado de conocimiento del catalán. Errores estándar en letra pequeña

Censo	1991				1996			
	Hombres		Mujeres		Hombres		Mujeres	
	Sabe	No sabe	Sabe	No sabe	Sabe	No sabe	Sabe	No sabe
LEER Y HABLAR								
Real	17.62	8.93	13.03	7.66	11.50	4.23	12.1	14.63
Pred:Saben1	9.21	8.38	15.06	5.65	13.83	4.18	17.03	4.77
	25.28	17.40	22.93	14.34	22.96	13.69	24.61	14.37
Pred:Nosaben	21.06	9.22	18.06	6.83	15.31	4.57	20.79	5.70
	26.60	18.50	25.22	16.02	24.19	14.44	27.16	15.92
Eff. D. Medio	-1.85	-0.84	-3.00	-1.18	-1.48	-0.39	-3.76	-0.94
	1.84	1.36	3.31	2.14	1.86	0.97	4.07	2.06
ESCRIBIR								
Real	21.19	10.47	15.09	8.06	14.03	5.86	14.32	5.41
Pred:Saben	23.35	10.09	18.63	6.68	17.08	6.12	20.19	6.49
	26.63	19.16	24.68	15.63	24.74	16.08	25.75	16.52
Pred:Nosaben	24.86	10.87	21.45	7.92	18.06	6.49	22.76	7.32
	27.47	20.04	26.41	17.18	25.40	16.64	27.17	17.65
Eff. D. Medio	-1.50	-0.78	-2.82	-1.24	-0.99	-0.37	-2.57	-0.83
	1.24	1.13	2.63	2.03	1.04	0.74	2.42	1.52

Tabla 5. Sobre-educación. Efectos discretos medios (en %) por nivel de conocimiento del catalán. Errores estándar en letra pequeña

Censo Sexo Conocimiento	1991				1996			
	Hombres		Mujeres		Hombres		Mujeres	
	Sabe	No sabe	Sabe	No sabe	Sabe	No sabe	Sabe	No sabe
LEER Y HABLAR PROBIT	-4.45	-2.05	-5.89	-2.19	-3.38	-0.99	-6.57	-1.68
	3.27	2.75	5.01	3.24	3.62	2.30	6.25	3.80
PROBIT BIVARIADO	-5.16	-2.02	-6.03	-1.78	-4.08	-0.92	-7.65	-1.44
	4.02	2.80	5.58	2.84	4.60	2.23	7.88	3.53
PROBIT + Ocu + Act	-1.63	-0.85	-2.55	-1.20	-1.27	-0.42	-3.61	-1.20
	1.56	1.36	2.70	2.15	1.53	1.03	3.67	2.51
BIVARIADO + Ocu + Act	-1.85	-0.84	-3.00	-1.18	-1.48	-0.39	-3.76	-0.94
	1.84	1.36	3.31	2.14	1.86	0.97	4.07	2.06
ESCRIBIR PROBIT	-4.81	-2.22	-6.10	-2.04	-2.87	-1.05	-4.89	-1.45
	2.88	2.69	4.34	2.94	2.59	1.94	4.03	2.79
PROBIT BIVARIADO	-4.72	-2.17	-5.92	-1.88	-2.87	-0.96	-5.25	-1.33
	2.88	2.58	4.46	2.65	2.71	1.79	4.70	2.64
PROBIT + Ocu + Act	-1.68	-0.87	-2.66	-1.20	-1.05	-0.42	-2.53	-0.94
	1.36	1.25	2.41	1.96	1.08	0.84	2.24	1.71
BIVARIADO + Ocu + Act	-1.50	-0.78	-2.82	-1.24	-0.99	-0.37	-2.57	-0.83
	1.24	1.13	2.63	2.03	1.04	0.74	2.42	1.52

Figura 1. Catalán hablado en Cataluña, País Valenciano y las Islas Baleares



Fuente: Vallverdú, 1990 y Censdata