

# ¿Cómo ha evolucionado la desigualdad de oportunidades educativas en España? Controlando el sesgo de selección de los modelos de transiciones educativas

*How has Inequality in Educational Opportunities Evolved in Spain?  
Controlling Selection Bias in Educational Transition Models*

**María Fernández-Mellizo**

## Palabras clave

Clase social

- Desigualdad de oportunidades educativas
- Expansión educativa
- Transiciones educativas

## Key words

Social Class

- Inequality of Educational Opportunities
- Educational Expansion
- Educational Transitions

## Resumen

En este artículo se analiza la evolución de la DOE a lo largo del siglo xx en España, incorporando más información sobre características personales y familiares de los estudiantes que en otros estudios. Esta aproximación permite controlar el posible sesgo de selección de los modelos de transiciones educativas, en los que muchos estudiantes van saliendo del análisis en las transiciones hacia niveles educativos superiores, haciendo que los resultados no sean extrapolables al conjunto de la población. Se ha utilizado una encuesta del CIS así como modelos de regresión logística. Controlando por dicho sesgo, se observa que la DOE permaneció constante, aunque también que la desigualdad disminuyó en el logro de la educación obligatoria entre las clases agrarias desde mitad de siglo, fruto posiblemente de la universalización de este nivel educativo.

## Abstract

This article analyses the evolution of inequality in educational opportunities over the 20<sup>th</sup> century in Spain, incorporating more information regarding students' personal and family characteristics than in other studies. This approach allows us to control for possible selection bias in educational transition models, in which many students are left out of analyses as transitions are made to higher levels of education, leading to results that cannot be extrapolated to the general population. A survey from Spain's Centre for Sociological Research (CIS) is used along with logistical regression models. Controlling for this bias, it is found that inequalities in educational opportunities have remained constant, although inequality in finishing obligatory education among agricultural classes decreased from the middle of the century, possibly the result of the universalisation of obligatory education.

## Cómo citar

Fernández-Mellizo, María (2022). «¿Cómo ha evolucionado la desigualdad de oportunidades educativas en España? Controlando el sesgo de selección de los modelos de transiciones educativas». *Revista Española de Investigaciones Sociológicas*, 177: 21-42. (doi: 10.5477/cis/reis.177.21)

La versión en inglés de este artículo puede consultarse en <http://reis.cis.es>

**María Fernández-Mellizo:** Universidad Complutense de Madrid | [mfmellizoso@edu.ucm.es](mailto:mfmellizoso@edu.ucm.es)

## MOTIVACIÓN Y PREGUNTAS DE LA INVESTIGACIÓN<sup>1</sup>

Este artículo tiene como objetivo analizar la evolución de la desigualdad de oportunidades educativas (a partir de ahora DOE) en España. El objeto de estudio no es en absoluto novedoso, puesto que ya se han realizado análisis similares por parte de otros investigadores<sup>2</sup>. Sin embargo, es llamativo que en dichos estudios no haya acuerdo sobre la evolución de la DOE a lo largo de tiempo, incluso cuando se analizan los mismos datos, algo que en parte refleja la heterogeneidad de sus planteamientos metodológicos. Además, en los análisis realizados del caso español no se ha controlado por la posible presencia de variables no observadas, con lo que los modelos utilizados adolecen de problemas derivados del sesgo de selección (Shavit y Blossfeld, 1993; Mare, 1993). Asimismo, la mayoría de los estudios que han analizado la evolución del impacto del origen social en el logro educativo para el caso español, en realidad todos los que incluyen el conjunto de transiciones educativas, solo llegan hasta cohortes nacidas a finales de los años sesenta (Fernández-Mellizo, 2014).

La presente investigación responde, pues, a esta doble necesidad de arrojar luz al debate, eminentemente metodológico pero, también sustantivo, sobre la evolución de la DOE en España y de actualizar los análisis existentes incorporando a personas más jóvenes, incluso aquellos nacidos en democracia. Para ello se ha analizado el es-

tudio n.º 2634 (Clases sociales y estructura social, de 2006) del CIS. Esta encuesta permite analizar la evolución de la DOE, estudiar el conjunto de transiciones educativas, llegar hasta cohortes bastante recientes (nacidas a principios de los años ochenta), e incorporar información sobre las notas de los estudiantes y la actitud de los padres hacia la educación, información que no suele estar presente en este tipo de encuestas. Solo se conoce un estudio parecido con esta encuesta (Martínez Celorrio, 2013), pero únicamente se refiere al logro del nivel universitario y la estadística empleada es descriptiva. Se ha seguido una metodología ortodoxa en los análisis de la evolución de la DOE (Shavit y Blossfeld, 1993) y tratado de controlar el sesgo de selección, al incorporar a los modelos más información sobre las características familiares y personales de los alumnos. De esta manera se trata, en primer lugar, de analizar cómo ha evolucionado la DOE en España a lo largo del siglo xx y, en segundo lugar, de observar si al controlar el sesgo de selección de los modelos los resultados se mantienen o no.

El artículo se estructura de la siguiente manera. En el segundo apartado se revisa brevemente la literatura, se desarrolla el argumento teórico y se plantean las hipótesis que han guiado la investigación. El tercer apartado aborda las cuestiones metodológicas: describe la encuesta utilizada, las variables empleadas, así como las técnicas de análisis de datos que se han usado. El cuarto apartado expone los resultados de la investigación en dos bloques: en primer lugar, se analiza la evolución de la DOE a través de modelos de transiciones educativas; en segundo lugar, se incorpora a los modelos información sobre la capacidad de los alumnos o las preferencias de los padres hacia la educación para tratar de disminuir su sesgo. En el último apartado se sintetizan los resultados poniéndolos en relación tanto con las hipótesis iniciales como con los estudios previos sobre el tema.

<sup>1</sup> Instituto de investigación TRANSOC de la UCM. Para esta investigación se ha contado con una Ayuda a la explotación del Banco de Datos del CIS concedida en 2016. Se agradecen los comentarios de Julio Carabaña y de tres revisores anónimos de la revista a una versión anterior de este artículo.

<sup>2</sup> Por citar algunas, las más completas: Peruga y Torres (1997); Carabaña (1999); Martínez (2002); Ballarino *et al.* (2009); Barone y Ruggera (2018). Véase una revisión sistemática de los estudios sobre la evolución de la DOE en España en Fernández-Mellizo (2014).

## ARGUMENTO TEÓRICO E HIPÓTESIS DE LA INVESTIGACIÓN

En la literatura internacional hay en la actualidad un debate sobre si durante el periodo de expansión de la educación de buena parte del siglo xx (desde los nacidos a principios de siglo hasta los nacidos a mediados de la década de los años setenta) la DOE se mantuvo constante o si, en cambio, disminuyó<sup>3</sup>. Desde la publicación de *Persistent Inequality* por Shavit y Blossfeld (1993) se cuestionó la tesis liberal (Blau y Duncan, 1967; Treiman, 1970) de que la expansión de la educación por sí misma igualaría las oportunidades educativas de los niños. Se constató, sin embargo, cómo en la mayoría de los trece países analizados, a la vez que el nivel educativo medio aumentaba, el diferencial entre los hijos de diferentes clases sociales permanecía constante. Posteriormente se ha desafiado este argumento (Breen *et al.*, 2009; Barone y Ruggera, 2018), aportando evidencia de una reducción de la DOE en un conjunto de países. La diferencia en los resultados encontrados se atribuye a diferencias metodológicas, fundamentalmente muestras más amplias (Erikson, 2019)<sup>4</sup>. En todo caso, de lo que no hay evidencia es de la eliminación de la desigualdad (Hadjar, 2019), y la disminución de la desigualdad sucede en un periodo concreto bastante excepcional, tras las Segunda Guerra Mundial (desde los nacidos en los años treinta a los nacidos a mitad de los cincuenta). Además, al considerar el valor en el mercado de trabajo de la educación la conclusión de que la DOE disminuyó puede no mantenerse para

todos los países (Erikson, 2019). También se han detectado algunos países en los que la DOE aumentó, aunque constituyen casos aislados con circunstancias muy excepcionales (Goldthorpe, 2007).

En España el grado de acuerdo sobre la evolución de la DOE no es mayor, incluso si cabe es menor, que en la literatura internacional (Fernández-Mellizo, 2014). La mayor parte de los estudios sobre este tema coinciden en describir un proceso de expansión de la educación desde los nacidos a principios de este siglo hasta los que lo hicieron a principios de la década de los años setenta. Sin embargo, el acuerdo ya no se produce ni siquiera al mostrar la evolución cuantitativa del logro educativo desde los nacidos a mediados de los años setenta (algunos describen la continuidad del proceso expansivo y, en cambio, otros muestran una mayor estabilidad en el nivel educativo medio de la población). El desacuerdo entre las diferentes investigaciones es total en cuanto al signo de la evolución de la DOE, incluso cuando se utilizan las mismas fuentes de datos. Por ejemplo, con los datos de la Encuesta Sociodemográfica del INE, que permite examinar la evolución de la DOE a través de varias transiciones educativas desde los nacidos a principios de siglo hasta los nacidos mediados de la década de los sesenta, algunos detectan subidas y bajadas (Peruga y Torres, 1997; Carabaña, 1999); otros, persistencia (Martínez García, 2002), mientras que incluso otros muestran una disminución del diferencial educativo por clases sociales (Ballarino *et al.*, 2009).

Al igual que en la literatura internacional, para el caso de España buena parte de las discrepancias en las tendencias detectadas se debe a la heterogeneidad metodológica de sus planteamientos. No solo se utilizan encuestas diferentes (las más comunes son la Sociodemográfica, la Encuesta de Población Activa —EPA— y la Encuesta de Condiciones de Vida —ECV—) o se operacionaliza la variable dependiente, el logro educativo, de forma distinta, sino que el pro-

<sup>3</sup> Las mejores revisiones de esta literatura internacional se encuentran en: Breen y Jonsson (2005), Goldthorpe (2007), Breen *et al.* (2009), Breen (2010), Shavit *et al.* (2007), Hadjar (2019) y Erikson (2019). Para una revisión en español véase también Fernández-Mellizo (2015). La literatura que se revisa se corresponde con las explicaciones racionales de la DOE, que tienen su origen en Boudon (1974).

<sup>4</sup> Además, en el libro de Shavit y Blossfeld (1993) se utiliza sobre todo *logit* binomial, mientras que Breen *et al.* (2009) y Barone y Ruggera (2018) utilizan *logit* ordinal.

pio método de estimación utilizado difiere entre los diversos estudios, alterando los resultados sustantivos sobre la evolución de la DOE (Fernández-Mellizo, 2014). Mientras que en algunos casos se utiliza la regresión lineal (Carabaña, 1999), en la mayoría se utilizan métodos no lineales, aunque en diferentes variantes: *logit* o *probit*, fundamentalmente.

La primera hipótesis de la investigación se formula de la siguiente manera:

H<sub>1</sub>: El diferencial de clase en el logro educativo se mantuvo amplio durante el siglo xx en España.

La constatada estabilidad de los diferenciales educativos por clase social ha recibido, por parte de la comunidad académica internacional, mucho mayor esfuerzo teórico que la posible variación en la DOE. Pese a su carácter *ad hoc*, teorías como la de la «desigualdad mantenida al máximo» (Raftery y Hout, 1993) o, con posterioridad, la «tesis de la pérdida de estatus» (Breen y Goldthorpe, 1997), constituyeron esfuerzos notables por dotar de cuerpo teórico a los patrones empíricos detectados. Por el contrario, el cambio en los diferenciales educativos por clase social apenas cuenta con justificación teórica, como los propios Breen *et al.* (2009) reconocen. La excepción más notable se encuentra, sin embargo, en Erikson y Jonsson (1996). En España, sin embargo, la mayoría de los análisis sobre la evolución de la DOE han sido eminentemente empíricos, sin una clara guía teórica o analítica. En los pocos casos en los que se recurre a la teoría se hace de manera *ad hoc* y utilizando teorías corroboradas en otros países.

Los modelos sobre la evolución de la DOE, llamados modelos de transiciones educativas (Shavit y Blossfeld, 1993; Mare, 1993), suelen tener problemas de sesgo de selección, derivado de la presencia de variables no observadas que afectan a la muestra de individuos analizados (aquellos que continúan estudiando) (Shavit y Blossfeld, 1993; Mare, 1981; Mare, 1993;

Cameron y Heckman, 1998; Shavit, Arum y Gamoran, 2007; Bernardi, 2012). Esto sucede porque a medida que los estudiantes atraviesan el sistema educativo van siendo seleccionados (o se autoseleccionan) en función de su capacidad, motivación y preferencias. Si tenemos modelos sesgados, las estimaciones obtenidas pueden no ser generalizables a la población de estudio, con lo que se cuestiona su validez interna. Una forma de controlar por la heterogeneidad no observada de sus análisis es incorporar más información sobre características familiares y personales. Otras estrategias utilizadas para disminuir este sesgo son la utilización de datos de hermanos para incorporar características familiares a los modelos (Mare, 1993) o la utilización de técnicas estadísticas que previenen este tipo de sesgo (Cameron y Heckman, 1998<sup>5</sup>; Bernardi, 2012).

Por tanto, introducir en los modelos información sobre la capacidad y motivación de los alumnos (sobre sus notas), así como sobre las preferencias de sus padres sobre la educación, puede reducir esta heterogeneidad no observada en los modelos de evolución de la DOE y construir de esta manera modelos menos sesgados. Llegamos de esta manera a la segunda, y última, hipótesis de la investigación:

H<sub>2</sub>: Los resultados del análisis de la evolución de la DOE cambian al controlar por las preferencias de los padres y la capacidad del alumno.

## METODOLOGÍA: DATOS, VARIABLES Y TÉCNICAS DE ANÁLISIS

Los datos utilizados para el análisis proceden, como ya se ha dicho, del estudio n.º 2634 (Clases sociales y estructura social, de 2006) del CIS. El análisis, por tanto, se

<sup>5</sup> Modelos ordinales de elección discreta.

realiza con una única encuesta de carácter transversal que contiene información sobre el año de nacimiento de las personas y, de esta manera, se pueden establecer cohortes de nacimiento que dan medida de la evolución en el tiempo. Esta estrategia analítica, de hecho, es la más utilizada en los estudios tanto dentro como fuera de España. Además de contener las preguntas relevantes para la investigación, el único requisito que han de cumplir estas encuestas es que deben tener un número de casos suficiente para lograr un mínimo de cohortes de nacimiento y transiciones educativas.

La encuesta del CIS tiene un tamaño muestral total de 8.265 casos. Sin embargo, no se pueden utilizar todos los casos para el análisis. La encuesta se ha realizado en 2006 a la población residente de 16 o más años (nacida en 1990 o antes). Hay que limitar los casos a aquellos que tengan como mínimo 25 años o, lo que es lo mismo, a aquellos que hayan nacido en 1981 o antes (si no, se corre el riesgo de que los que tengan de 16 a 24 años no hayan estudiado más simplemente porque no les ha dado tiempo). Al aplicar este límite a los casos la muestra se reduce un 14% (unos 1.150 casos), pero hace el análisis más preciso. El tamaño muestral, aunque no muy amplio para el tipo de análisis planteado, resulta suficiente minimizando adecuadamente el número de cohortes de nacimiento y transiciones educativas utilizadas. Barone y Ruggera (2018) excluyen de su análisis a los países que no lleguen a los 2.000 casos. Si bien existen encuestas más amplias con las que se suelen hacer este tipo de análisis, como la Sociodemográfica, la ECV o la EPA del INE, la del CIS es la única que incorpora variables de capacidad y motivación de los estudiantes y de las preferencias de los padres sobre la escolarización, de ahí que se haya empleado para esta investigación.

Las variables utilizadas en el análisis (resumidas en las tablas 3 y 4 del Anexo) son las siguientes. La variable dependiente es

el nivel educativo alcanzado, la más utilizada entre los sociólogos de estratificación actuales. Se han considerado cuatro categorías: tener menos de estudios básicos obligatorios (T0); titular en educación obligatoria general (T1)<sup>6</sup>; titular en educación secundaria posobligatoria o FP (T2)<sup>7</sup>; titular en educación universitaria (T3). Esta clasificación, tal cual, refleja el máximo nivel educativo alcanzado por el individuo. Hay que notar que es un proceso acumulativo; es decir, en T2 están los que tienen T1 más los que tienen T2, en T3 los que tienen T1 más T2 más T3 (de hecho, por ejemplo, T3 puede aumentar sin que haya un porcentaje mayor de gente que, teniendo T2, haya decidido ir a T3, simplemente porque T2, o T1, hayan aumentado). Para que el proceso deje de ser acumulativo se han considerado las transiciones condicionadas a haber conseguido el nivel educativo previo (que, en realidad, constituyen variables diferentes porque el número de casos varía): T2/T1 se refiere a conseguir T2 cuando se ha conseguido ya T1; T3/T2 se refiere a conseguir T3 cuando se ha conseguido ya T2. Esta forma de concebir el logro educativo, como una decisión educativa secuencial, es la dominante en los estudios sobre la evolución de la DOE desde los trabajos de Mare (1980, 1981).

Se han utilizado cohortes de nacimiento para medir el paso del tiempo y así examinar la evolución de la DOE; en particular, se han construido cuatro: 1908-1949, 1950-1959, 1960-1969 y 1970-1981. No se han creado más cohortes (idealmente deberían ser de 10 en 10 años) debido al ta-

<sup>6</sup> Se refiere a terminar con éxito la ESO, la EGB o el Bachillerato elemental, en función de cuándo se ha estudiado. Se excluye la FP I del sistema LGE. La codificación del CIS agrupa la FP de grado medio del sistema LOGSE con la FP I de la LGE, con lo que, aunque tenga sentido analítico separarlas, no es posible.

<sup>7</sup> Incluye toda la FP, incluida la que no es educación secundaria posobligatoria, como FP I del sistema LGE o FP de grado superior de la LOGSE.

maño muestral y, en todo momento, hay que tener en cuenta que la primera cohorte engloba un periodo de tiempo de más de 40 años, superior al resto de cohortes, que solo incluyen 10 años. La utilización de cohortes de nacimiento es el método estándar en este tipo de estudios.

Respecto a la variable independiente principal, la clase social de origen, se ha construido una clasificación adaptando al esquema EGP (Erikson-Goldthorpe-Portocarero)<sup>8</sup> a España y al tamaño muestral (que no permite muchas categorías). Este esquema de clases weberiano se construye a partir de la ocupación del individuo, y es el más utilizado entre los sociólogos de estratificación social, tanto nacionales como internacionales. Se han construido seis clases sociales a partir de la información sobre la ocupación del padre cuando el hijo tenía 16 años: directivos, profesionales y técnicos (I+II); trabajadores no manuales (IIIab); empresarios y autónomos (IVab); trabajadores manuales cualificados (V+VI); trabajadores manuales no cualificados (VIIa); agricultores y peones agrarios (IVc+VIIb)<sup>9</sup>.

Otras variables independientes utilizadas han sido la capacidad del alumno y las preferencias de sus padres respecto al estudio. La capacidad del alumno se ha medido a través de una variable subjetiva de las notas que sacaba a los 16 años (o antes si se abandonaron los estudios) en cinco valores; el menor si se sacaban muy malas notas y el mayor si se sacaban muy buenas notas. Las preferencias de los padres se han incorporado a través de una variable que mide la actitud de los padres

hacia el estudio a los 16 años (o antes si hay abandono previo) en tres categorías: negativa, si los padres inducían a los hijos a abandonar cuanto antes los estudios para empezar a trabajar; indiferente, cuando a los padres no les importaba si sus hijos estudiaban o cuánto estudiaban; y positiva, cuando los padres animaban a sus hijos a estudiar más y a continuar con los estudios. Estas variables no suelen estar presentes en este tipo de análisis (debido a que no se tiene normalmente información sobre las mismas).

Como variables de control se han utilizado dos que suelen estar presentes en este tipo de estudios: el género, en formato estándar, y el país en el que se ha vivido la mayor parte del tiempo hasta cumplir los 16 años (se ha diferenciado entre los que vivieron en España y en otro país).

Como método de estimación de los parámetros se ha utilizado la regresión logística, siguiendo el «modelo de transiciones educativas» de Mare que distingue los cambios en las tasas de escolarización de la asociación entre el origen social del alumno y el logro educativo, en consonancia con los análisis realizados en el libro de Shavit y Blossfeld (1993)<sup>10</sup>. Al igual que estos estudios, la utilización de los modelos *logit* responde a una decisión previa sobre cómo interpretar los parámetros; se utilizan los modelos *logit* porque se interpretan los parámetros en términos de *odds ratio*, consistente con la definición sociológica de desigualdad de oportunidades educativas. Los modelos *logit* y *probit* hacen supuestos diferentes sobre la distribución del error, logístico para el *logit* y normal para el *probit*. De hecho, ambos modelos presentan resultados muy similares, casi iguales en muestras como las que se utilizan en ciencias sociales (Fernández-Mellizo, 2015).

<sup>8</sup> Véase Erikson, Goldthorpe y Portocarero (1979), así como Erikson y Goldthorpe (1992).

<sup>9</sup> No se ha podido distinguir entre empresarios grandes y pequeños, y se han puesto todos juntos (salvo los del sector agrario) en una categoría. Agradezco enormemente la ayuda de Miguel Caínzos para la construcción de la variable de clase social de origen.

<sup>10</sup> Del mismo modo permiten controlar los cambios en la estructura ocupacional a lo largo del tiempo, sin duda muy relevantes en España.

## RESULTADOS DE LA INVESTIGACIÓN

Los resultados del análisis se presentan en dos apartados: en primer lugar, se analiza la evolución de la DOE en España durante el siglo xx, contextualizado en un periodo de expansión de la escolarización; en segundo lugar, se trata de corregir el sesgo de selección de los modelos introduciendo determinadas características familiares y personales de los alumnos.

### EVOLUCIÓN DE LA DESIGUALDAD DE OPORTUNIDADES EDUCATIVAS EN ESPAÑA

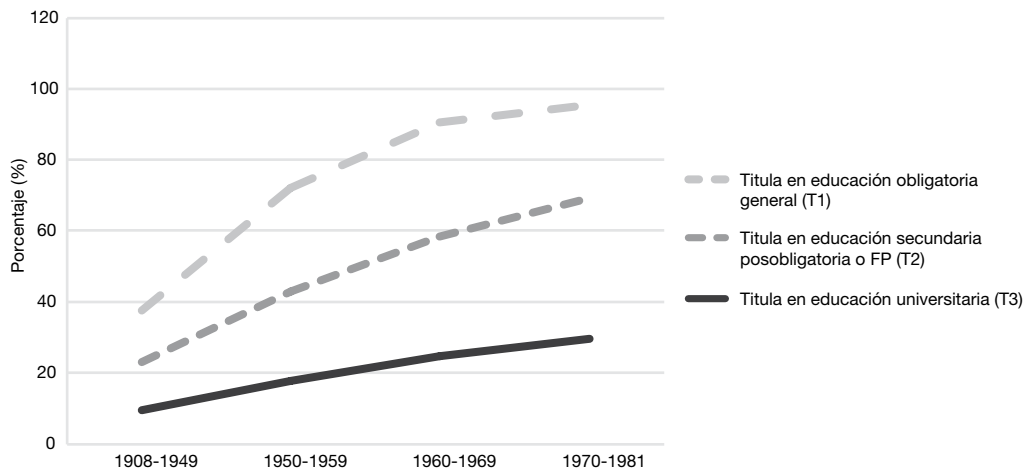
Antes de presentar la evidencia de la evolución de la DOE, merece la pena detenerse un poco en el contexto de expansión de la educación. Hay numerosa evidencia internacional de la expansión de la educación en el siglo xx (Haim y Shavit, 2013). Un estudio de 33 países (sobre todo de Europa) concluye que la educación se expandió de forma sostenida en todo el periodo (Hadjari y Becker, 2016). Para España también existe evidencia de expansión educativa en este periodo de tiempo, desde los nacidos a principios de siglo hasta al menos los nacidos en la década de los años setenta (Fernández-Mellizo, 2014). El aumento del nivel educativo de la población española resulta evidente a la luz de las estadísticas históricas de España de la Fundación BBVA (Núñez, 2005) y otras que utilizan datos más recientes (Fuente y Doménech, 2016). Las discrepancias a este respecto son menores; algunos encuentran que la expansión continúa hasta los nacidos a principios de los ochenta (Paolo, 2012), mientras otros describen un estancamiento desde los nacidos a finales de los años setenta (Carabaña, 2013).

Desde un punto de vista sociológico, importa más el título educativo que se consigue que el número de años que se estudia,

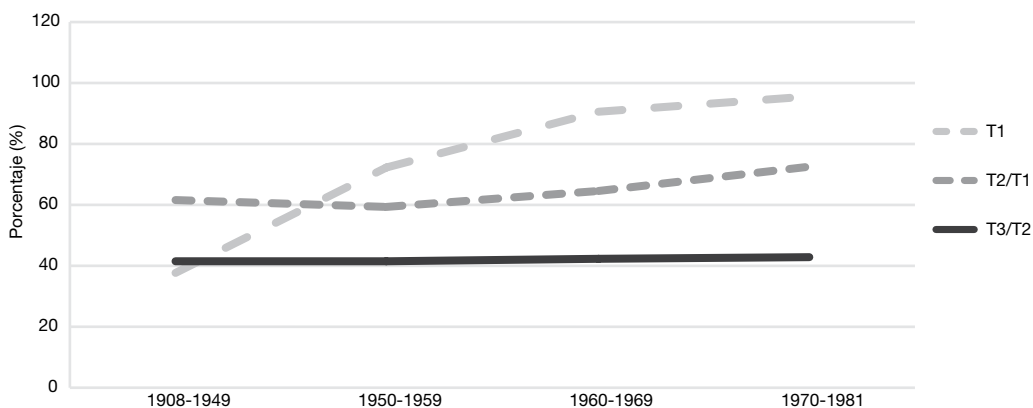
porque el título es lo que permite la inserción laboral y, a la postre, va a determinar la posición social de esa persona. Por ello, a continuación hacemos un análisis de transiciones educativas, tanto del máximo nivel educativo alcanzado como de las transiciones educativas condicionadas (a haber alcanzado el nivel educativo anterior). Como se ha explicado en el apartado anterior, los análisis estadísticos se han realizado con cuatro cohortes debido a que la encuesta no es lo suficientemente amplia para analizar un número mayor de cohortes. De cara a la comparación entre las cohortes, hay que tener en cuenta que la primera cohorte engloba a los nacidos en un periodo de tiempo muy superior (cuatro décadas) a las otras cohortes (más o menos una década).

El gráfico 1 muestra el porcentaje de personas de cada cohorte que ha alcanzado cada nivel educativo. Por razones de tamaño muestral previamente explicadas se han considerado tres niveles educativos: titular en educación obligatoria general (T1); titular en educación secundaria posobligatoria o FP (T2); titular en educación universitaria (T3).

Atendiendo por tanto a los títulos educativos, la expansión de la educación en España entre los nacidos en el siglo xx es clara. Mientras que entre los nacidos en la primera mitad del siglo solo titulaba en educación general obligatoria el 38%, en educación secundaria posobligatoria o FP el 23%, y en educación universitaria el 10%, en la generación más joven, entre los nacidos a partir de los años setenta, el 96% alcanza el nivel obligatorio general (llegando prácticamente a la saturación de ese nivel educativo), el 69% las enseñanzas medias (incluida toda la FP) y el 30% el nivel universitario. En todos los niveles educativos ha habido aumento, aunque es fácil advertir en el gráfico cómo a medida que se asciende de nivel educativo el incremento es menor; es decir, la expansión ha sido más rápida en los niveles educativos bajos que en los niveles educativos altos.

**GRÁFICO 1.** *Porcentaje de personas de cada cohorte que alcanza un determinado nivel educativo*

Fuente: CIS estudio n.º 2634.

**GRÁFICO 2.** *Porcentaje que alcanza un nivel educativo dado que tiene el anterior (decisión secuencial) por cohorte*

Notas: T1: Titular en educación obligatoria general; T2: Titular en educación secundaria posobligatoria o FP; T3: Titular en educación universitaria; T2/T1: Conseguir T2 cuando se ha conseguido ya T1; T3/T2: Conseguir T3 cuando se ha conseguido ya T2.

Fuente: CIS estudio n.º 2634.

El gráfico 2 presenta una información complementaria a la anterior, la transición de nivel educativo condicionada a haber obtenido el nivel anterior. Puesto que para cursar un determinado nivel educativo hay que haber completado con éxito el anterior, interesa saber qué porcentaje de los que han completado un determinado nivel educativo consiguen

el siguiente. Eso nos da una idea más precisa de lo que se expande cada nivel educativo en sí mismo, no por efecto de la expansión en un nivel educativo previo. Por ejemplo, para cursar la educación secundaria posobligatoria o la FP hay que haber superado con éxito la educación general obligatoria; interesa en este caso saber qué porcentaje de los que

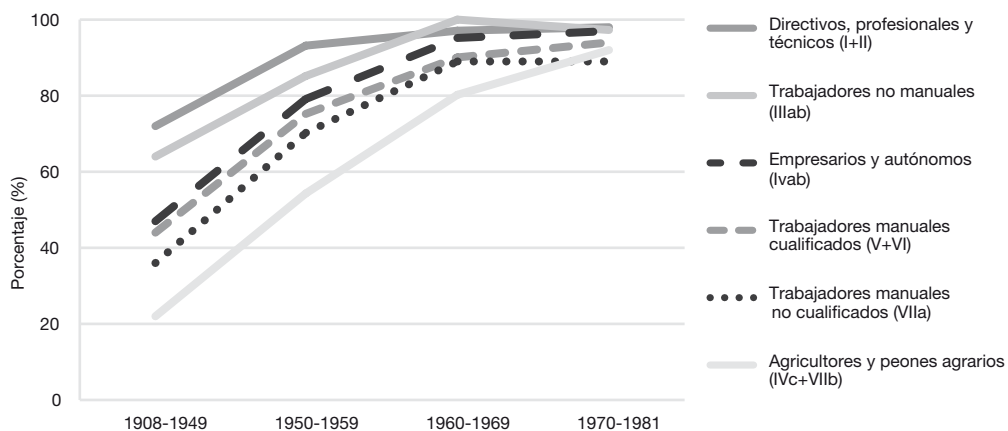


han completado la educación general obligatoria continúa (y acaba con éxito) la educación secundaria posobligatoria o la FP.

El porcentaje de los que titulan en educación obligatoria general es igual que en el gráfico anterior. Sin embargo, los porcentajes de los que hacen las transiciones condicionadas son más altos que los de las transiciones sin condicionar, puesto que aquellos que no superan el nivel previo salen de los cálculos. Por ejemplo, entre los nacidos a partir de 1970, el 30% de la cohorte consigue el título de educación universitaria, mientras que este porcentaje asciende al 43% si solo se tiene en cuenta a los que han superado el nivel educativo previo, es decir, han titulado en educación secundaria posobligatoria o FP. Se excluye, por tanto, de los cálculos a aquellos que no han conseguido el nivel educativo previo, T2, puesto que es imposible que cursen T3, lo que aumenta lógicamente el porcentaje.

El aumento entre las generaciones es muy acusado para el caso de la educación obligatoria general, tal y como se refleja en el gráfico 2. Sin embargo, las transiciones educativas condicionadas (a haber alcanzado el nivel educativo anterior) muestran una mayor estabilidad a lo largo del tiempo que en el caso de las transiciones sin establecer condiciones. El porcentaje de los que titulan en educación secundaria posobligatoria o FP (habiendo completado el nivel anterior) se mantiene estable entre los nacidos en la primera parte del siglo xx, en torno al 60%; en cohortes posteriores este porcentaje aumenta hasta el 73% entre los nacidos a partir de la década de los setenta. Más llamativa es aún la estabilidad del porcentaje de los que titulan en educación universitaria (habiendo completado el nivel anterior), que se mantiene en torno al 42% en todo el periodo considerado.

**GRÁFICO 3.** *Porcentaje con educación general obligatoria (T1) por clase social de origen y cohorte*



Fuente: CIS estudio n.º 2634.

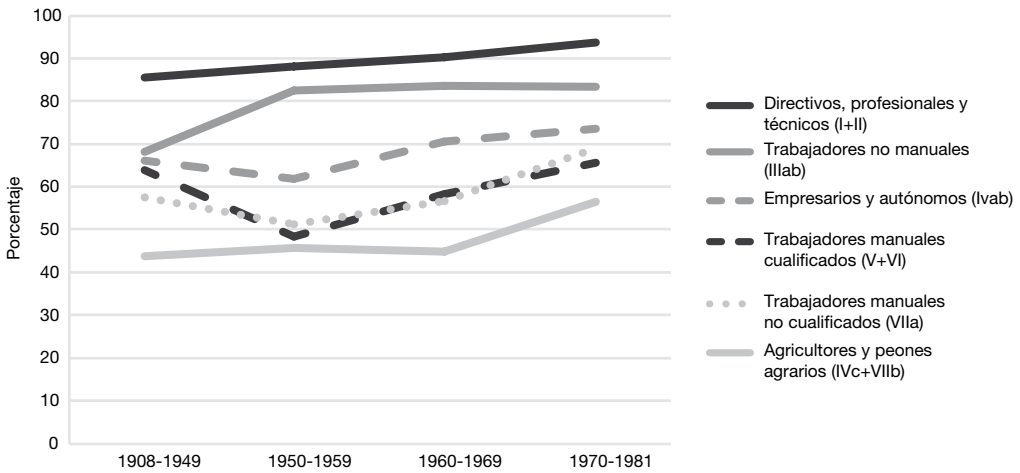
Para analizar la evolución de la desigualdad de oportunidades educativas en España se ha introducido en el análisis la variable de clase social de origen, a partir de la información sobre la ocupación del padre cuando el hijo tenía 16 años. El análisis de la evolución

de la DOE muestra el papel del origen social en los modelos de transiciones educativas utilizando técnicas no lineales, como el *logit*, para obtener los parámetros estimados. Se examina el impacto de la clase social en la obtención de títulos educativos y, de forma

más precisa, en las transiciones de niveles educativos condicionadas a haber obtenido el nivel educativo previo, como se hace de forma estándar en la literatura internacional. Para comenzar este análisis, se ha realizado un análisis descriptivo de la evolución por cohortes de nacimiento del porcentaje de cada clase social que logra cada nivel educativo. El

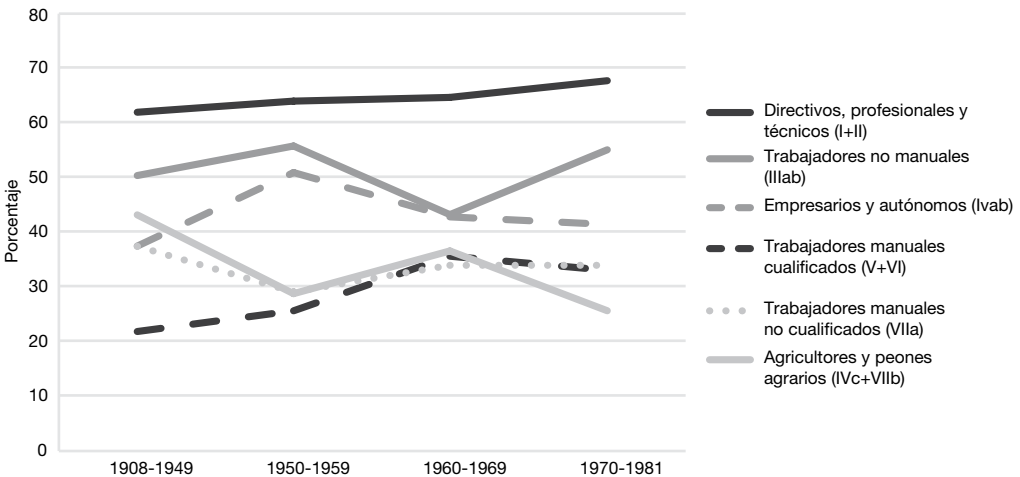
gráfico 3 muestra el porcentaje de cada clase social que obtiene el título de educación obligatoria general, así como su evolución a través de las cohortes de nacimiento. Se aprecia una evidente reducción de las diferencias entre las clases sociales, derivado posiblemente de la práctica saturación de las clases más altas, en las últimas cohortes de nacimiento.

**GRÁFICO 4.** *Porcentaje con educación posobligatoria o FP que ya ha obtenido educación general obligatoria (T2/T1) por clase social y cohorte*



Fuente: CIS estudio n.º 2634.

**GRÁFICO 5.** *Porcentaje con educación universitaria que ya ha obtenido educación posobligatoria o FP (T3/T2) por clase social y cohorte*



Fuente: CIS estudio n.º 2634.

Sin embargo, en la primera de las transiciones condicionadas, obtener educación secundaria posobligatoria o FP tras haber completado la educación obligatoria general (T2/T1), apenas se aprecia una muy ligera reducción de las diferencias entre las clases sociales, tal y como se observa en el gráfico 4. En la segunda de las transiciones condicionadas, el porcentaje que obtiene educación universitaria dado que ya ha obtenido previamente educación secundaria posobligatoria o FP (T3/T2), la evolución de las diferencias depende de cada clase en cuestión y es difícil establecer una conclusión general, tal y como se muestra en el gráfico 5.

En general, de los anteriores gráficos se desprende que la reducción de las diferencias entre clases sociales en la obtención de títulos educativos no resulta evidente, salvo en el caso del nivel educativo más bajo (la obtención de la educación obligatoria general, T1). La obtención de la educación general obligatoria se ha ampliado mucho durante el periodo de tiempo analizado y las clases sociales más altas, como ya se ha comentado, tienen una representación muy alta desde los años cincuenta y, sobre todo, sesenta. La reducción de la desigualdad que se observa es precisamente debido a que las clases más altas ya no tienen mucho margen para mejorar.

Para hacer un tratamiento estadístico más adecuado de este tipo de datos, se van a hacer modelos de regresión logística ya que, como ha sido explicado en la metodología, la variable dependiente es dicotómica. Para simplificar el análisis, siguiendo la literatura internacional sobre este tema, se van a presentar solo las transiciones condicionadas (T2/T1 y T3/T2) además de T1. Se incluye tanto la variable de las cohortes de nacimiento, siendo la base de comparación el primer periodo, como la de clase social, cuya base de comparación son los directivos y profesionales. También se incluyen, como variables de control, el género y el país en el que ha vivido la mayor parte de los años transcurridos hasta los 16 años.

Para ver cómo evoluciona la DOE para cada clase social se incluyen las interacciones de las diferentes cohortes con cada clase social. Como muestra la tabla 1, en relación a T1, la expansión educativa en este nivel educativo durante todo el periodo resulta evidente, tal y como muestran los coeficientes positivos (significativos) de las cohortes de nacimiento; la base de la comparación es la cohorte más antigua así que un coeficiente positivo indica que las subsiguientes generaciones presentan un logro educativo mayor. Los hombres y los que han vivido fuera de España hasta los 16 años tienen también más probabilidad de realizar la transición. La variable de género está medida como una *dummy* con valor 1 si se es hombre, con lo que el coeficiente positivo (significativo) indica más probabilidad de realizar esta transición en el caso de los hombres. Por su parte, la variable de país es otra *dummy* con valor 1 si la mayor parte del tiempo transcurrido hasta los 16 años se ha pasado en España, con lo que el signo negativo (significativo) indica que es más probable que los que no han vivido en España realicen esa transición. El impacto de la clase social también es evidente, tal y como muestran los coeficientes negativos (significativos), salvo para los hijos de los trabajadores no manuales, cuyo coeficiente no es significativo. Los signos negativos son lógicos puesto que se compara el logro de cada clase con el de los hijos de directivos y profesionales; si el logro es menor el signo es, por tanto, negativo. Por último, la estabilidad de la DOE en esta transición es clara, al no resultar significativos los coeficientes de ninguna de las interacciones entre la cohorte de nacimiento y la clase social. Respecto a T2/T1, la expansión de la educación solo se produjo para la última cohorte, a partir de los nacidos en los años setenta, y la variable género deja de ser significativa. Además, aunque la tónica dominante es la estabilidad de la DOE (las interacciones entre las cohortes y el origen social no salen

significativas), se aprecia el aumento de la desigualdad entre los trabajadores manuales cualificados a partir de los años setenta (único coeficiente significativo). Esto último se observa en el coeficiente negativo, que indica que el logro de los hijos de los trabajadores manuales cualificados en comparación con los hijos de directivos y profesionales se ha reducido desde los nacidos a principios de siglo hasta los nacidos en los años setenta. En lo que concierne a T3/T2, ya no hay evidencia de expansión educativa, y también el país de nacimiento deja

de importar (ningún coeficiente es significativo). Se vuelve a apreciar la estabilidad de la DOE durante todo el periodo, salvo para el sector agrícola a partir de los años setenta (entre los que aumenta la desigualdad). Todos los coeficientes menos este último salen no significativos. Así, el coeficiente negativo (significativo) de los hijos de los trabajadores agrícolas nacidos en los años setenta indica que ha aumentado la diferencia de logro educativo en comparación con los hijos de los directivos y profesionales nacidos a principios de siglo.

**TABLA 1.** Modelos de regresión logística para estimar el logro de diferentes niveles educativos

Variables	T1	T2/T1	T3/T2
<i>Cohortes (base: nacidos 1908-1949)</i>			
Nacidos 1950-1959	1,650***	0,080	0,100
Nacidos 1960-1969	2,490***	0,320	0,140
Nacidos 1970-1981	3,030***	0,810**	0,250
Género	0,360***	0,070	0,000
País	-0,480**	-1,030***	0,170
<i>Clase social (base: directivos y profesionales)</i>			
Trab. no manuales	-0,400	-1,110***	-0,420
Empresarios y autónomos	-1,120***	-1,160***	-1,000***
Trab. manuales cualif.	-1,220***	-1,260***	-1,770***
Trab. manuales no cualif.	-1,530***	-1,590***	-1,060***
Sector agrícola	-2,290***	-2,060***	-0,770**
<i>Interacción año de nacimiento y clase social (base: nacidos en 1908-1949; directivos y profesionales)</i>			
Nacidos en 1950-1959; trab. no manuales	-0,490	0,760	0,050
Nacidos en 1950-1959; empresarios y autónomos	-0,200	-0,270	0,450
Nacidos en 1950-1959; trab. manuales cualif.	-0,340	-0,740	0,110
Nacidos en 1950-1959; trab. manuales no cualif.	-0,270	-0,240	-0,420
Nacidos en 1950-1959; sector agrícola	-0,190	-0,070	-0,690
Nacidos en 1960-1969; trab. no manuales		0,570	-0,480
Nacidos en 1960-1969; empresarios y autónomos	0,480	-0,160	0,070
Nacidos en 1960-1969; trab. manuales cualif.	-0,040	-0,580	0,540
Nacidos en 1960-1969; trab. manuales no cualif.	0,150	-0,340	-0,260
Nacidos en 1960-1969; sector agrícola	0,190	-0,340	-0,390
Nacidos en 1970-1981; trab. no manuales	-0,040	0,010	-0,100
Nacidos en 1970-1981; empresarios y autónomos	0,440	-0,530	-0,070
Nacidos en 1970-1981; trab. manuales cualif.	-0,030	-0,790*	0,310
Nacidos en 1970-1981; trab. manuales no cualif.	0,530	-0,320	-0,350
Nacidos en 1970-1981; sector agrícola	0,607	-0,400	-1,020**
Constante	1,290***	2,790***	0,310
N	5.674	4.217	2.769
Pseudo-R <sup>2</sup>	0,320	0,080	0,050

Nota: \* p < 0,10; \*\* p < 0,05; \*\*\* p < 0,01.

Fuente: CIS estudio n.º 2634.

Como se puede apreciar en la anterior tabla, el proceso de expansión educativa es claro en T1, pero solo se produce para los nacidos a partir de la década de los setenta para T2/T3 e, incluso, no se detecta expansión educativa alguna para T3/T2. Además, características como el género o el país en el que se ha vivido los primeros años pierden importancia a medida que aumenta el nivel educativo. Respecto a la influencia de la clase social, su inclusión en el modelo aumenta el ajuste del modelo en todas las transiciones (aumenta el pseudo- $R^2$ )<sup>11</sup> y su impacto es el esperado. Además, se observa cómo el impacto de la clase social es diferente en función del nivel educativo (aumenta del paso de T1 a T2/T1 y disminuye en T3/T2, indicando que la principal selección social se produce en T2). En relación a la evolución de la desigualdad de oportunidades educativas no se encuentra evidencia de que haya habido alguna variación significativa, dado que las interacciones con las cohortes de nacimiento no salen significativas (con la excepción de un aumento de la desigualdad entre los nacidos a partir de los años setenta entre los trabajadores manuales cualificados para T2/T3 y en el sector agrícola para T3/T2). Por tanto, los resultados del análisis de transiciones educativas apuntan a la estabilidad de la desigualdad de oportunidades educativas<sup>12</sup>.

<sup>11</sup> Las tablas en las que se incluyen las variables de una en una no se han mostrado aquí por razones de espacio. Si el lector desea consultarlas puede solicitarlas.

<sup>12</sup> Se ha probado a hacer modelos en los que la cohorte de referencia sea la segunda (los nacidos en la década de los años cincuenta), sin detectar cambios significativos en los resultados. La razón de esta simulación es que en el análisis cohorte a cohorte, no mostrado aquí, se observa que la influencia de la clase social en la probabilidad de conseguir T1 se incrementa hasta los nacidos a finales de los años cincuenta y luego se reduce hasta los nacidos a principios de los años sesenta (véase la tabla 5 del Anexo).

## **TRATANDO DE CONTROLAR LA HETEROGENEIDAD NO OBSERVADA DE LOS MODELOS DE TRANSICIONES EDUCATIVAS**

Tanto la capacidad o motivación de los alumnos como las preferencias de sus padres hacia la educación pueden tener un impacto en la probabilidad de realizar las transiciones educativas. No incorporar información sobre estas variables, tal y como se hace en la literatura internacional y como se ha hecho en el apartado anterior, puede sesgar los resultados del análisis de la evolución de la DOE al centrarse el análisis en un grupo homogéneo no representativo del conjunto de la población de estudio (los que siguen estudiando). Según reconocen los propios Shavit y Blossfeld (1993), lo mejor sería contar con más información para poder controlar la heterogeneidad no observada de los modelos. De hecho, Mare (1993) utiliza una estrategia parecida a la aquí mostrada (derivada de información sobre hermanos, no de preguntas directas como la que disponemos). Dado que la encuesta del CIS nos permite controlar por estas variables, se van a replicar los análisis del apartado anterior con dichos controles, para ver si se mantienen, o no, los resultados. La tabla 2 muestra los resultados de este ejercicio.

En primer lugar, se observa que tanto las notas del estudiante como la actitud hacia el estudio por parte de los padres están relacionadas positivamente con la probabilidad de realizar cada una de las transiciones educativas. Los coeficientes positivos (significativos) indican que cuanto mayores son las notas, mayor el logro académico, y que si los padres tienen una actitud indiferente o positiva hacia la educación (en comparación con una actitud negativa) el nivel educativo alcanzado es mayor. La única excepción la encontramos en T3/T2 para los casos en los que los padres tienen una actitud indiferente hacia la escolarización, frente a aquellos que tienen

una actitud negativa (el coeficiente, positivo, no es significativo). En segundo lugar, la introducción de nueva información, las notas del estudiante y las preferencias de sus padres hacia la educación, mejoran la capacidad predictiva de todos los modelos (la pseudo-R<sup>2</sup> aumenta), sobre todo en T2/T1 y T3/T2.

Finalmente, la introducción de estas nuevas variables en los modelos disminuye el impacto de la clase social en las probabilidades de transición educativa (los coeficientes disminuyen)<sup>13</sup>. La clase social estaba recogiendo, en los modelos sin controlar, efectos de otras variables omitidas.

**TABLA 2.** Modelos de regresión logística para estimar el logro educativo controlando por heterogeneidad no observada

Variables	T1	T2/T1	T3/T2
<i>Cohortes (base: nacidos 1908-1949)</i>			
Nacidos 1950-1959	2,070***	0,290	0,110
Nacidos 1960-1969	2,690***	0,520	0,080
Nacidos 1970-1981	2,890***	1,000**	0,220
Género	0,470***	0,250***	0,110
País	-0,300	-0,960***	0,280**
<i>Clase social (base: directivos y profesionales)</i>			
Trab. no manuales	0,080	-0,770*	-0,340
Empresarios y autónomos	-0,080***	-0,880**	-0,920**
Trab. manuales cualif.	-0,770***	-0,700*	-1,560***
Trab. manuales no cualif.	-0,880***	-1,000**	-0,690
Sector agrícola	-1,710***	-1,600***	-0,670*
Notas	0,610***	0,880***	0,780***
<i>Actitud de los padres hacia el estudio (base: negativa)</i>			
Indiferente	0,340***	0,440**	0,110
Positiva	1,200***	1,400***	1,310***
<i>Interacción año de nacimiento y clase social (base: nacidos en 1908-1949; directivos y profesionales)</i>			
Nacidos en 1950-1959; trab. no manuales	-1,320*	0,460	0,020
Nacidos en 1950-1959; empresarios y autónomos	-0,680	-0,320	0,390
Nacidos en 1950-1959; trab. manuales cualif.	-0,810	-1,180**	-0,070
Nacidos en 1950-1959; trab. manuales no cualif.	-0,750	-0,510	0,720
Nacidos en 1950-1959; sector agrícola	-0,600	-0,080	-0,640
Nacidos en 1960-1969; trab. no manuales		0,600	-0,350
Nacidos en 1960-1969; empresarios y autónomos	0,170	-0,380	-0,000
Nacidos en 1960-1969; trab. manuales cualif.	-0,290	-0,910*	0,360
Nacidos en 1960-1969; trab. manuales no cualif.	-0,090	-0,680	-0,500
Nacidos en 1960-1969; sector agrícola	-0,270	-0,530	-0,330
Nacidos en 1970-1981; trab. no manuales	0,310	-0,010	0,000
Nacidos en 1970-1981; empresarios y autónomos	0,550	-0,730	-0,120
Nacidos en 1970-1981; trab. manuales cualif.	-0,110	-1,100**	0,160
Nacidos en 1970-1981; trab. manuales no cualif.	-0,370	-0,650	-0,640
Nacidos en 1970-1981; sector agrícola	0,480	-0,640	-1,030**
Constante	-2,120***	-2,020	-4,120***
N	5.213	4.0770	2.712
Pseudo-R <sup>2</sup>	0,360	0,190	0,130

Nota: \* p < 0,10; \*\* p < 0,05; \*\*\* p < 0,01.

Fuente: CIS estudio n.º 2634.

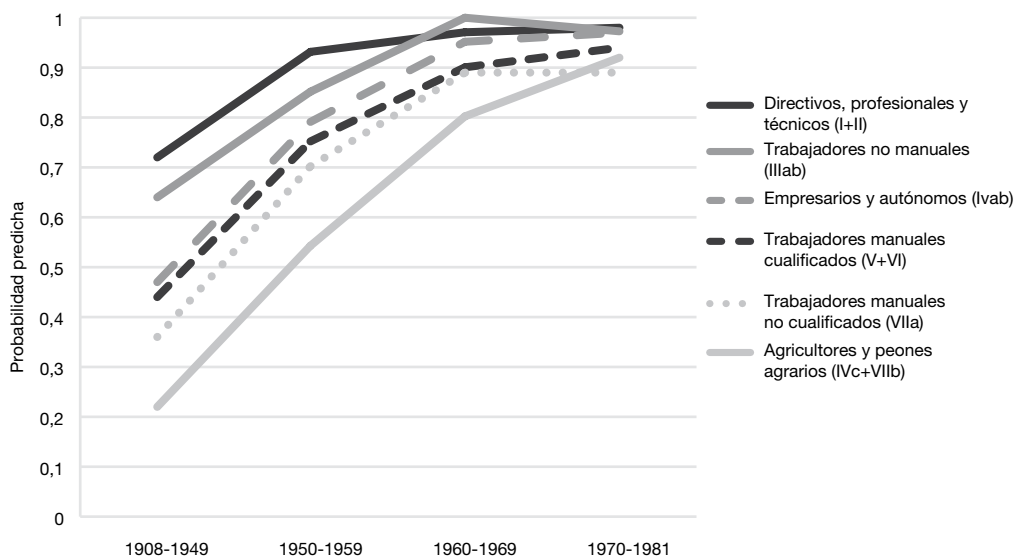
<sup>13</sup> Los resultados detallados no se han mostrado aquí por razones de espacio. Si lo desea puede solicitarlos.

Si comparamos esta tabla con la mostrada en el apartado anterior (tabla 1), sin controlar por esta nueva información, vemos que los resultados del análisis de la evolución de la DOE son muy similares; la abrumadora mayoría de interacciones de la clase social con las cohortes no son significativas, con lo que de nuevo hay evidencia favorable a la hipótesis de la estabilidad de la DOE. Se observan, sin embargo, algunas interacciones aisladas significativas; además de los ya detectados previamente (aumentos de desigualdad en T2/T1 para los trabajadores manuales cualificados y en T3/T2 para los del sector agrícola a partir de los nacidos en la década de los setenta), se observan aumentos también en T1 para los

trabajadores no manuales nacidos en la década de los años cincuenta y en T2/T1 para los trabajadores manuales cualificados de cualquier cohorte de edad (los coeficientes negativos —significativos— muestran, como se ha explicado anteriormente, el menor logro de esa categoría en relación a los directivos y profesionales nacidos en la primera mitad de siglo).

De modo colateral se observa también que al introducir la nueva información el género es significativo en T2/T1, ser hombre aumenta la probabilidad de titular educación secundaria, y el país de nacimiento en T3/T2, haber nacido en España aumenta la probabilidad de conseguir un título de educación universitaria.

**GRÁFICO 6.** Probabilidad de obtener educación obligatoria según clase social por cohorte controlando por heterogeneidad no observada (segunda cohorte como base de comparación)



Fuente: CIS estudio n.º 2634.

Al igual que en el apartado anterior, se han replicado los análisis poniendo la segunda cohorte (los nacidos en los años cincuenta) como cohorte de referencia. Además, se han hecho análisis introduciendo por separado la capacidad (notas) y las preferencias de los padres hacia la

educación. La única diferencia encontrada con los análisis aquí mostrados es que se encuentra, en los modelos que solo incluyen la variable de las notas y cuya cohorte de referencia es la segunda, una interacción significativa para T1 que muestra una reducción de la desigualdad en ese

nivel para los del sector agrícola nacidos a partir de la década de los años setenta (en relación a los nacidos en los años cincuenta), tal y como indica el signo positivo del coeficiente. Los resultados del análisis se encuentran en la tabla 6 del Anexo. Para interpretar mejor los resultados de la regresión logística, en el gráfico 6 se presentan los valores predichos marginales (al 95% de intervalo de confianza). Estos valores representan la probabilidad predicha de obtener educación obligatoria según clase social de origen y cohorte de nacimiento.

Por tanto, al introducir nueva información en los modelos, los resultados no cambian sustantivamente; las tendencias de la desigualdad de oportunidades educativas se mantienen fundamentalmente constantes, al igual que lo hacían en los modelos sin controlar. La única alteración relevante, producto de incorporar las notas del estudiante, es que se encuentra la disminución de la desigualdad entre los agricultores y peones agrícolas (en relación a los directivos y profesionales) nacidos entre los años cincuenta y setenta del siglo xx. El sesgo derivado de los modelos de transiciones educativas, de no incluir dicha información, no es, por tanto, tan abultado como indica la literatura consultada (Shavit y Blossfeld, 1993; Mare, 1981; Mare, 1993; Cameron y Heckman, 1998; Shavit, 2007; Bernardi, 2012). En todo caso, con ánimo de ser prudentes, siempre existe la posibilidad de que haya otra información relevante que no se haya tenido en cuenta.

## CONCLUSIONES

A lo largo de estas páginas se han presentado los resultados de la investigación sobre la evolución de la DOE en España utilizando datos del CIS, que permiten llegar hasta cohortes más jóvenes y que dan la posibilidad de incluir más información para

obtener resultados menos sesgados. Sabemos que durante el siglo xx el nivel educativo de los españoles aumentó, al igual que lo hizo en el resto de países de su entorno. Ahora bien, la expansión educativa fue muy intensa en el nivel educativo más bajo (educación obligatoria y general), menor en niveles intermedios (educación secundaria posobligatoria y FP) y casi inexistente en el nivel universitario.

El grueso de la investigación corresponde a medir la evolución de la DOE en España. Conviene recordar que el número de casos de la encuesta utilizada no es muy amplio, y que la metodología que se usa tiene sus limitaciones. Aún así, los análisis realizados apuntan a la estabilidad de la DOE en el periodo de tiempo considerado, puesto que solo se detectan algunos cambios muy puntuales en la DOE, en particular subidas de alguna clase social en alguna cohorte aislada (aumentos de desigualdad para los nacidos en la década de los setenta tanto en T2/T1 para los trabajadores manuales cualificados como en T3/T2 para los del sector agrícola). La expansión de la educación, por tanto, no habría conducido en España a la disminución de la desigualdad educativa por clase social; el nivel educativo medio de la población aumentó, pero, sin embargo, las desigualdades por origen social se mantuvieron en todo el periodo, tal y como postulamos en la primera hipótesis. Esta evidencia es consistente con lo hallado para España, con datos de la Encuesta Sociodemográfica, por Martínez García (2002), pero no así con lo encontrado por otros autores; Ballarino *et al.* (2009) detectaron una disminución de la DOE, así como Peruga y Torres (1997) y Carabaña (1999) aumentos y disminuciones de la DOE. Fuera de España, los resultados de esta investigación avalan las conclusiones de Shavit y Blossfeld (1993) sobre la persistencia de los diferenciales de clase social en el logro educativo a través del tiempo y, en cambio, no van en la línea



de las conclusiones de Breen *et al.* (2009) y Barone y Ruggera (2018), que mostraron una disminución de esos diferenciales. Así pues, se confirma la evidencia que apunta a que, en España, el proceso de expansión educativa experimentado por los nacidos en buena parte del siglo xx (hasta los nacidos en los años ochenta, en democracia) no fue acompañado por una disminución de la desigualdad en el logro educativo por origen social, sino que dicho proceso de elevación general del nivel educativo convivió con la persistencia o mantenimiento de las diferencias educativas entre aquellos de diferentes clases sociales.

La parte más novedosa de la investigación es la que se refiere al control de la heterogeneidad no observada de los modelos del análisis de la evolución de la DOE. De hecho, es esta parte la que justifica el análisis con los datos de la encuesta de CIS, que permite introducir en los modelos información sobre las preferencias de los padres y la capacidad de los alumnos. La incorporación de estas nuevas variables, ausentes en otras investigaciones, nos ha permitido tratar de reducir el posible sesgo de selección (autoselección) de estos análisis de transiciones educativas (Shavit y Blossfeld, 1993; Mare, 1993), obteniendo parámetros más fiables. Los resultados del análisis corroboran las conclusiones previas sobre la persistencia de los diferenciales de clase en el logro educativo; solo se encuentran variaciones en la DOE para clases y cohortes aisladas, alejándose de la pauta general de estabilidad (se observan aumentos de desigualdad en T1 para los trabajadores no manuales nacidos en la década de los años cincuenta, aumentos de desigualdad en T2/T1 para los trabajadores manuales cualificados de cualquier cohorte de edad, y en T3/T2 para los del sector agrícola nacidos a partir de la década de los setenta). No obstante, aunque esa sea la tónica general, se advierte que la inclusión de la información sobre las notas de los alumnos, reflejo de su

capacidad y motivación, modifica la predicción de los modelos. En particular, en los niveles bajos de educación, incorporar las notas de los alumnos hace que disminuya la desigualdad de oportunidades educativas entre los alumnos nacidos entre los años cincuenta y setenta con padres del sector agrario. Ello respalda la confirmación, aunque limitada, de la segunda de las hipótesis. Al incorporar esta información estamos disminuyendo el sesgo de selección que inevitablemente produce el sistema educativo al irse quedando con los mejores estudiantes y, de ese modo, obtenemos modelos menos sesgados. No se puede descartar, sin embargo, que la incorporación de otras variables omitidas al modelo reduzca aún más dicho sesgo. En la medida en que nuevas encuestas vayan incorporando más información se podrán conseguir modelos mejor especificados y más fiables.

## BIBLIOGRAFÍA

- Ballarino, Gabriele; Bernardi, Fabrizio; Requena, Miguel y Schadee, Hans (2009). «Persistent Inequalities? Expansion of Education and Class Inequality in Italy and Spain». *European Sociological Review*, 25(1): 123-138. doi: 10.1093/esr/jcn031
- Barone, Carlo y Ruggera, Lucia (2018). «Educational Equalization Stalled? Trends in Inequality of Educational Opportunity between 1930 and 1980 across 26 European Nations». *European Societies*, 20: 1-25. doi: 10.1080/14616696.2017.1290265
- Bernardi, Fabrizio (2012). «Unequal Transitions: Selection Bias and the Compensatory Effect of Social Background in Educational Careers». *Research in Social Stratification and Mobility*, 30: 159-174. doi: 10.1016/j.rssm.2011.05.005
- Blau, Peter M. y Duncan, Otis D. (1967). *The American Occupational Structure*. New York: John Wiley and Sons.
- Boudon, Raymond (1974). *Education, Opportunity, and Social Inequality: Changing Prospects in Western Society*. New York: Wiley.
- Breen, Richard (2010). «Educational Expansion and Social Mobility in the 20th Century». *Social Forces*, 89(2): 365-388. doi: 10.1353/sof.2010.0076

- Breen, Richard y Goldthorpe, John H. (1997). «Explaining Educational Differentials. Towards a Formal Rational Action Theory». *Rationality and Society*, 9(3): 275-305. doi: 10.1177/104346397009003002
- Breen, Richard y Jonsson, Jan O. (2005). «Inequality of Opportunity in Comparative Perspective: Recent Research on Educational Attainment and Social Mobility». *Annual Review of Sociology*, 31: 223-243. doi: 10.1146/annurev.soc.31.041304.122232
- Breen, Richard; Luijckx, Ruud; Müller, Walter y Polak, Reinhard (2009). «Nonpersistent Inequality in Educational Attainment: Evidence from Eight European Countries». *American Journal of Sociology*, 114(5): 1475-1521. doi: 10.1086/595951
- Cameron, Stephen V. y Heckman, James J. (1998). «Life Cycle Schooling and Dynamic Selection Bias: Models and Evidence for Five Cohorts of American». *Journal of Political Economy*, 106: 262-334. doi: 10.1086/250010
- Carabaña, Julio (1999). *Dos Estudios sobre Movilidad Intergeneracional*. Madrid: Fundación Argentina.
- Carabaña, Julio (2013). «Crecimiento del bachillerato e igualdad de los años ochenta». *Revista de la Asociación de Sociología de la Educación*, 6(1): 6-31. Disponible en: <https://ojs.uv.es/index.php/RASE/article/view/8595>, acceso 15 de enero de 2021.
- Di Paolo, Antonio (2012). «Parental Education and Family Characteristics: Educational Opportunities across Cohorts in Italy and Spain». *Revista de Economía Aplicada*, 58(XX): 119-146. Disponible en: <https://www.redalyc.org/pdf/969/96924442005.pdf>, acceso 15 de enero de 2021.
- Erikson, Richard (2019). «How Does Education Depend on Social Origin?». En: Becker, R. (ed.). *Research Handbook on the Sociology of Education*. Massachusetts: E-Elgar.
- Erikson, Robert y Goldthorpe, John H. (1992). *The Constant Flux: A Study of Class Mobility in Industrial Societies*. Oxford: Oxford UP.
- Erikson, Robert y Jonsson, Jan O. (1996). *Can Education be Equalized? The Swedish Case in Comparative Perspective*. Boulder, Colorado: Westview.
- Erikson, Robert; Goldthorpe, John H. y Portocarero, Lucienne (1979). «Intergenerational Class Mobility in Three Western European Societies: England, France and Sweden». *The British Journal of Sociology*, 30: 415-441. doi: 10.2307/589632
- Fernández-Mellizo, María (2014). «La evolución de la desigualdad de oportunidades educativas: una revisión sistemática de los análisis del caso español»/«The Evolution of Inequality of Educational Opportunities: A Systematic Review of Analyses of the Spanish Case». *Revista Española de Investigaciones Sociológicas*, 147: 107-120. doi: 10.5477/cis/reis.147.107
- Fernández-Mellizo, María (2015). «Continuidad o cambio en la desigualdad de oportunidades educativas: evidencia internacional y teorías». *Revista Española de Sociología*, 23: 151-164. Disponible en: <https://recyt.fecyt.es/index.php/res/article/view/65378>, acceso 15 de enero de 2021.
- Fuente, Ángel de la y Doménech, Rafael (2016). «El nivel educativo de la población en España y sus regiones: 1960-2011». *Investigaciones Regionales*, 34: 73-94. Disponible en: <https://investigacionesregionales.org/es/article/el-nivel-educativo-de-la-poblacion-en-espana-y-sus-regiones-1960-2011/>, acceso 15 de enero de 2021.
- Goldthorpe, John H. (2007). *On Sociology: Numbers, Narratives, and the Integration of Research and Theory*. Oxford: Oxford University Press (2.ª ed.).
- Hadjar, Andreas (2019). «Educational Expansion and Inequalities: How Did Inequalities by Social Origin and Gender Decrease in Modern Industrial Societies?». En: Becker, R. (ed.). *Research Handbook on the Sociology of Education*. Massachusetts: E-Elgar.
- Hadjar, Andreas y Becker, Rolf (2016). «Education Systems and Meritocracy: Social Origin, Educational and Status Attainment». En: Hadjar, A. y Gross, C. (eds.). *Education systems and inequalities*. Bristol: Policy Press.
- Haim, Eyal B. y Shavit, Yossi (2013). «Expansion and Inequality of Educational Opportunity: A Comparative Study». *Research in Social Stratification and Mobility*, 31: 22-31. doi: 10.1016/j.rssm.2012.10.001
- Mare, Robert D. (1980). «Social Background and School Continuation Decisions». *Journal of the American Statistical Association*, 75(370): 295-305. doi: 10.1080/01621459.1980.10477466
- Mare, Robert D. (1981). «Change and Stability in Educational Stratification». *American Sociological Review*, 46: 72-87. doi: 10.2307/2095027
- Mare, Robert D. (1993). «Educational Stratification on Observed and Unobserved Components of Family Background». En: Shavit, Y. y Blossfeld, H. *Persistent Inequality. Changing Educational Attainment in Thirteen Countries*. Boulder: Westview.

- Martínez Celorrio, Xavier (2013). «Tendencias de movilidad y reproducción social por la educación en España». *Revista de la Asociación de Sociología de la Educación*, 6(1): 32-41. Disponible en: <https://ojs.uv.es/index.php/RASE/article/view/8596>, acceso 15 de enero de 2021.
- Martínez García, José S. (2002). *Habitus o calculus? Dos intentos de explicar la desigualdad de oportunidades educativas de los nacidos en España entre 1907 y 1966, con datos de la Encuesta Socio-Demográfica*. Carabaña, Julio (dir.), Madrid: Universidad Autónoma de Madrid. [Tesis doctoral].
- Núñez, Clara Eugenia (2005). «Educación». En: Carreras, A. y Tafunell, X. (coords.). *Estadísticas históricas de España. Siglos XIX-XX*. Bilbao: Fundación BBVA.
- Peruga, Rodrigo y Torres Mora, José A. (1997). «Desigualdad educativa en la España del siglo XX: un estudio empírico». En: VV. AA. *Educación, vivienda e igualdad de oportunidades. II Simposio sobre Igualdad y Distribución de la Renta y la Riqueza*. Madrid: Fundación Argenteria-Visor.
- Raftery, Adrian E. y Hout, Michael (1993). «Maximally Maintained Inequality: Expansion, Reform, and Opportunity in Irish Education, 1921-75». *Sociology of Education*, 66(1): 41-62. doi: 10.2307/2112784
- Shavit, Yossi y Blossfeld, Hans-Peter (1993). *Persistent Inequality. Changing Educational Attainment in Thirteen Countries*. Boulder: Westview.
- Shavit, Yossi; Arum, Richard y Gamoran, Adam (2007). *Stratification in Higher Education: A Comparative Study*. Palo Alto, California: Stanford University Press.
- Treiman, Donald J. (1970). «Industrialization and Social Stratification». *Sociological Inquiry*, 40: 207-234. doi: 10.1111/j.1475-682X.1970.tb01009.x

**RECEPCIÓN:** 18/11/2019

**REVISIÓN:** 05/05/2020

**APROBACIÓN:** 23/11/2020

## ANEXO

TABLA 3. Descriptivos variables cuantitativas y cualitativas ordinales

Variable	Tipo de variable	Observaciones	Media	Desviación típica	Mínimo	Máximo
Número de años de estudio	Cuantitativa	7.338	11,58	4,59	0	25
Notas	Cualitativa ordinal	7.306	3,58	0,84	1 (muy malas)	5 (muy buenas)
Año de nacimiento	Cuantitativa	7.096	1.956,78	16,59	1.908	1.981

Fuente: CIS estudio n.º 2634.

TABLA 4. Descriptivos variables cualitativas nominales

Variable	Tipo	Observaciones	Valores	Frecuencias
Género	Cualitativa nominal dicotómica	8.265	0: Mujer	2.917
			1: Hombre	3.106
Actitud de los padres hacia el estudio	Cualitativa nominal	7.222	1: Negativa	827
			2: Indiferencia	1.158
			3: Positiva	5.236
País en el que vivió hasta los 16 años	Cualitativa nominal dicotómica	8.239	0: Otro país	580
			1: España	8.239
Clase social de origen	Cualitativa nominal	7.510	1: Directivos y profesionales	806
			2: Trab. no manuales	518
			3: Empresarios y autónomos	1.272
			4: Trab. manuales cualificados	1.754
			5: Trab. manuales no cualificados	1.095
			6: Sector agrícola	2.061
Cohortes de nacimiento	Cualitativa nominal	7.088	1: 1908-1949	2.376
			2: 1950-1959	1.159
			3: 1960-1969	1.504
			4: 1970-1981	2.046
Nivel educativo máximo alcanzado	Cualitativa nominal	7.502	1: Menos de estudios básicos obligatorios (T0)	1.755
			2: Estudios básicos obligatorios y generales (T1)	2.039
			3: Estudios posobligatorios no universitarios y de FP (T2)	2.251
			4: Estudios universitarios (T3)	1.455
Alcanzar T2 frente a alcanzar solo T1 (T2/T1)	Cualitativa nominal dicotómica	5.751	0: Estudios básicos obligatorios y generales (T1)	2.040
			1: Estudios postobligatorios no universitarios y de FP (T2)	3.710
Alcanzar T3 frente a alcanzar solo T2 (T3/T2)	Cualitativa nominal dicotómica	3.691	0: Estudios posobligatorios no universitarios y de FP (T2)	2.241
			1: Estudios universitarios (T3)	1.449

Fuente: CIS estudio n.º 2634.

**TABLA 5.** Modelos de regresión logística para estimar T1 «cohorte a cohorte»

VARIABLES	1908-1949	1950-1959	1960-1969	1970-1981
Género	0,490***	0,550***	0,310*	-0,660***
País	-2,180***	-0,920**	-1,190**	0,960***
<i>Clase social (base: directivos y profesionales)</i>				
Trab. no manuales	-0,380	-0,850		-0,440
Empresarios y autónomos	-1,100***	-1,310***	-0,620	-0,650
Trab. manuales cualif.	-1,210***	-1,540***	-1,240**	-1,340**
Trab. manuales no cualif.	-1,530***	-1,770***	-1,360**	-0,950
Sector agrícola	-2,270***	-2,480***	-2,080***	-1,580***
Constante	2,890***	3,270***	4,480***	3,680
N	1.616	984,000	1.209	1.795
Pseudo-R <sup>2</sup>	0,100	0,080	0,060	0,050

Nota: \* p < 0,10; \*\* p < 0,05; \*\*\* p < 0,01.

Fuente: CIS estudio n.º 2634.

**TABLA 6.** Modelos de regresión logística para estimar T1 controlando por heterogeneidad no observada (segunda cohorte como base de comparación)

VARIABLES	M1	M2	M3
<i>Cohortes (base: nacidos 1950-1959)</i>			
Nacidos 1908-1949	-1,400***	-1,450***	-2,060***
Nacidos 1960-1969	1,270***	1,240***	0,670
Nacidos 1970-1981	1,970***	2,000***	0,960
Género	0,370***	0,460***	0,450***
País	-0,510**	-0,280	-0,250
<i>Clase social (base: directivos y profesionales)</i>			
Trab. no manuales	-0,380	-0,180	-1,150*
Empresarios y autónomos	-1,060***	-1,090***	-1,68***
Trab. manuales cualif.	-1,330***	-1,270***	-1,840***
Trab. manuales no cualif.	-1,510***	-1,300***	-1,880***
Sector agrícola	-2,240***	-2,130***	-2,660***
Notas		0,700***	0,700***
<i>Interacción año de nacimiento y clase social (base: nacidos en 1950-59; directivos y profesionales)</i>			
Nacidos en 1908-1949; trab. no manuales			1,030
Nacidos en 1908-1949; empresarios y autónomos			0,560
Nacidos en 1908-1949; trab. manuales cualif.			0,750
Nacidos en 1908-1949; trab. manuales no cualif.			0,570
Nacidos en 1908-1949; sector agrícola			0,540
Nacidos en 1960-1969; trab. no manuales			
Nacidos en 1960-1969; empresarios y autónomos			0,840
Nacidos en 1960-1969; trab. manuales cualif.			0,520
Nacidos en 1960-1969; trab. manuales no cualif.			0,590
Nacidos en 1960-1969; sector agrícola			0,420

**TABLA 6.** Modelos de regresión logística para estimar T1 controlando por heterogeneidad no observada (segunda cohorte como base de comparación) (Continuación)

Variables	M1	M2	M3
Nacidos en 1970-1981; trab. no manuales			1,140
Nacidos en 1970-1981; empresarios y autónomos			1,190
Nacidos en 1970-1981; trab. manuales cualif.			0,780
Nacidos en 1970-1981; trab. manuales no cualif.			1,240
Nacidos en 1970-1981; sector agrícola			1,280*
Constante	2,710***	0,050	0,580
N	5.779	5.436	5.331
Pseudo-R <sup>2</sup>	0,320	0,340	0,330

Nota: \* p < 0,10; \*\* p < 0,05; \*\*\* p < 0,01.

Fuente: CIS estudio n.º 2634.

# How has Inequality in Educational Opportunities Evolved in Spain? Controlling Selection Bias in Educational Transition Models

*¿Cómo ha evolucionado la desigualdad de oportunidades educativas en España?  
Controlando el sesgo de selección de los modelos de transiciones educativas*

**María Fernández-Mellizo**

## Key words

Social Class  

- Inequality of Educational Opportunities
- Educational Expansion
- Educational Transitions

## Palabras clave

Clase social  

- Desigualdad de oportunidades educativas
- Expansión educativa
- Transiciones educativas

## Abstract

This article analyses the evolution of inequality in educational opportunities over the 20<sup>th</sup> century in Spain, incorporating more information regarding students' personal and family characteristics than in other studies. This approach allows us to control for possible selection bias in educational transition models, in which many students are left out of analyses as transitions are made to higher levels of education, leading to results that cannot be extrapolated to the general population. A survey from Spain's Centre for Sociological Research (CIS) is used along with logistical regression models. Controlling for this bias, it is found that inequalities in educational opportunities have remained constant, although inequality in finishing obligatory education among agricultural classes decreased from the middle of the century, possibly the result of the universalisation of obligatory education.

## Resumen

En este artículo se analiza la evolución de la DOE a lo largo del siglo xx en España, incorporando más información sobre características personales y familiares de los estudiantes que en otros estudios. Esta aproximación permite controlar el posible sesgo de selección de los modelos de transiciones educativas, en los que muchos estudiantes van saliendo del análisis en las transiciones hacia niveles educativos superiores, haciendo que los resultados no sean extrapolables al conjunto de la población. Se ha utilizado una encuesta del CIS así como modelos de regresión logística. Controlando por dicho sesgo, se observa que la DOE permaneció constante, aunque también que la desigualdad disminuyó en el logro de la educación obligatoria entre las clases agrarias desde mitad de siglo, fruto posiblemente de la universalización de este nivel educativo.

## Citation

Fernández-Mellizo, María (2022). "How has Inequality in Educational Opportunities Evolved in Spain? Controlling Selection Bias in Educational Transition Models". *Revista Española de Investigaciones Sociológicas*, 177: 21-42. (doi: 10.5477/cis/reis.177.21)

**María Fernández-Mellizo:** Universidad Complutense de Madrid | mfmellizoso@edu.ucm.es

## AIM AND RESEARCH QUESTIONS<sup>1</sup>

The aim of this article is to analyse the evolution of inequality in educational opportunities (IEO) in Spain. This objective is by no means new, as similar studies have been carried out by other researchers<sup>2</sup>. However, it is striking that the findings from these studies on the evolution of IEO differ, even when analysing the same data, which in part reflects a heterogeneity in their methodological approaches. In addition, in the analyses carried out in the case of Spain, the possible presence of unobserved variables has not been controlled for; thus, the models used suffer from problems related to selection bias (Shavit and Blossfeld, 1993; Mare, 1993). In addition, the majority of studies that have analysed the evolution of the impact of social origin on educational achievement in Spain—in reality all those that include all the educational transitions—only reach cohorts born up until the end of the 1960s (Fernández-Mellizo, 2014).

This study is a response to this dual necessity of shedding light on both the methodological and substantive debate, as well as on the evolution of IEO in Spain, and to update existing analyses by incorporating younger cohorts into the analyses, including those born under Spanish democracy. To do this we analyse survey no. 2.634 (Social classes and social structure, 2006) of Spain's Centre for Sociological Research (CIS). This survey allows us to analyse the evolution of IEO, study the totality of educational transitions, include recent cohorts (those born at the beginning of the 1980s), and incorpo-

rate information on students' grades and parental attitudes toward education, information which is generally not available in these types of surveys. We know of only one similar study using this survey (Martínez Celorrio, 2013), which only refers to reaching university and the statistic employed is descriptive. We follow an orthodox methodology in our analyses of the evolution of IEO (Shavit and Blossfeld, 1993) and we attempt to control for selection bias by incorporating more information regarding the family and personal characteristics of the students in the models. In this way, we not only analyse the evolution of IEO in Spain over the 20th century but examine if the results vary when controlling for the selection bias of the models.

The article is structured in the following manner. In the second section we briefly review the literature and develop our theoretical argument and present the hypotheses that have guided the research. The third section addresses methodological questions: describing the survey used, the variables and the techniques of data analysis employed. The fourth section presents the results in two blocks: first, we analyse the evolution of IEO through educational transition models; secondly, we incorporate information regarding the ability of the students and parents' preferences regarding education into the models to attempt to decrease their bias. In the last section we synthesise the results, relating them to both our initial hypotheses and to earlier studies.

## THEORETICAL ARGUMENT AND HYPOTHESES OF THE STUDY

Currently, a debate exists in the international literature over whether inequality in educational opportunities remained constant or if, in contrast, it declined during the expansion of education that took place over a good part of the 20<sup>th</sup> century (from those born at the beginning of the century to the middle of

<sup>1</sup> TRANSOC research institute of the UCM. This research has been carried out with the support of a CIS Data Bank grant awarded in 2016. The author would like to thank Julio Carabaña and three anonymous reviewers of this journal for their comments regarding an earlier version of this article.

<sup>2</sup> To cite some, the most complete are: Peruga and Torres (1997); Carabaña (1999); Martínez (2002); Ballarino *et al.* (2009); Barone and Ruggera (2018). For a systematic review of studies on IEO in Spain see Fernández-Mellizo (2014).



the 1960s)<sup>3</sup>. Since the publication of *Persistent Inequality* by Shavit and Blossfeld (1993), the liberal thesis that the expansion of education by itself equalised children's educational opportunities (Blau and Duncan, 1967; Treiman, 1970) has come under question. The authors found that in the majority of the thirteen countries they analysed that while the average level of education increased, the differential between the children of different social classes remained constant. Subsequently, this finding has been challenged (Breen *et al.*, 2009; Barone and Ruggera, 2018), providing evidence of a decline in IEO in a number of countries. The difference in the results found is attributed to methodological differences, fundamentally larger samples (Erikson, 2019)<sup>4</sup>. In any case, there is no evidence of the elimination of inequality (Hadjar, 2019), and the decline in inequality took place in a period which is quite exceptional, after the Second World War (from those born in the 1930s to those born in the middle of the 1950s). In addition, in considering the value of education in the labour market, the conclusion that IEO declined cannot be maintained for all the countries (Erickson, 2019). It has also been found that IEO increased in some countries, although they constitute isolated cases under very exceptional circumstances (Goldthorpe, 2007).

The level of agreement found among studies regarding the evolution of IEO in Spain is not higher, and may even be lower, than found in the international literature (Fernández-Mellizo, 2014). The majority of the studies on this issue describe a process of

the expansion of education from the beginning of the 20<sup>th</sup> century until the decade of the 1970s. However, there is no agreement regarding the evolution of educational achievement of those born in the mid-1970s (some describe a continuity of the expansionary process while, in contrast, others find greater stability in the average educational level of the population). The disagreement among different studies is complete in terms of the direction of the evolution of IEO, even when the same source of data is used. For example, with data from the INE Sociodemographic Survey, which permits researchers to look at the evolution of IEO through various educational transitions for those born at the beginning of the century up through those born in the middle of the 1960s, some detect increases or decreases (Peruga and Torres, 1997; Carabaña, 1999), others persistence (Martínez García, 2002), while others show a decline in educational differences by social class (Ballarino *et al.*, 2009).

As with the international literature, for the Spanish case many of the differences in the trends detected are due to the methodological heterogeneity of the approaches. Not only do researchers use different surveys (the most common ones are the INE's Sociodemographic Survey, Labour Force Survey and Living Conditions Survey) or operationalise the dependent variable of educational achievement differently, but the estimation method used differs among the different studies, altering the substantive results regarding the evolution of IEO (Fernández-Mellizo, 2014). While in some cases linear regression is employed (Carabaña, 1999), in the majority of studies non-linear methods are used, although with different variations: fundamentally *logit* or *probit*.

The first hypothesis of this study is the following:

H<sub>1</sub>: Differences in educational attainment by social class in Spain in the 20<sup>th</sup> century remained significantly large.

<sup>3</sup> The best reviews of this international literature are found in: Breen and Jonsson (2005), Goldthorpe (2007), Breen *et al.* (2009), Breen (2010), Shavit *et al.* (2007), Hadjar (2019) and Erikson (2019). For a review in Spanish see Fernández-Mellizo (2015). The literature reviewed corresponds with rational explanations for IEO, which have their origin in Boudon (1974).

<sup>4</sup> In addition, in the book by Shavit and Blossfeld (1993) they used binomial logistic regression, while Breen *et al.* (2009) and Barone and Ruggera used ordinal logistic regression.

The observed stability of educational differences by social class have received much greater theoretical attention from the international academic community than possible variations in IEO. Despite their ad hoc character, theories such as that of “maximum maintained inequality” (Raftery and Hout, 1993) or, subsequently, Breen and Goldthorpe’s thesis of a “loss of status” (1997), constituted notable efforts to provide theoretical foundations to the empirical patterns found. In contrast, the change in educational differentials by social class are barely provided with any theoretical justification, as Breen *et al.* (2009) recognise. The most notable exception is found in Erikson and Jonsson (1996). In Spain, however, the majority of analyses of the evolution of IEO have been empirical, without a clear theoretical or analytical framework. In the few cases in which an appeal to theory is made, it is done in an *ad hoc* manner, using theories tested in other countries.

Models on the evolution of IEO, referred to as models of educational transitions (Shavit and Blossfeld, 1993; Mare, 1993), tend to have problems with selection bias resulting from the presence of unobserved variables that affect the sample of individuals analysed (those that continue studying) (Shavit and Blossfeld, 1993; Mare, 1981; Mare, 1993; Cameron and Heckman, 1998; Shavit, Arum and Gamoran, 2007; Bernardi, 2012). This happens because as students advance in the educational system, they are selected (or self-selected) in function of their ability, motivation and preferences. If we have biased models, the estimates obtained cannot be generalised to the study population, leading to questions of their internal validity. One way of controlling for unobserved heterogeneity in the analysis is to incorporate more information about family and personal characteristics. Other strategies used to decrease bias are the use of data on siblings to incorporate family characteristics in the models (Mare, 1993) or the

use of statistical techniques that prevent this type of bias (Cameron and Heckman, 1998<sup>5</sup>; Bernardi, 2012).

Therefore, introducing information about the ability and motivation of students (regarding their grades) into the models, as well as the preferences of their parents regarding education, can reduce this unobserved heterogeneity in the models of the evolution of IEO and help in the construction of less biased models. Considering this, we have our second and final research hypothesis:

H<sub>2</sub>: The results of the analysis of the evolution of IEO change when we control for parents’ preferences and students’ abilities.

## METHODOLOGY: DATA, VARIABLES AND ANALYSIS TECHNIQUES

The data used for the analysis come from study no. 2.634 (Social classes and social structure), a 2006 CIS survey. The analysis, therefore, is based on a single transversal survey that contains information regarding the year of birth of participants; in this way, it permits the establishment of birth cohorts that provide a measure of evolution over time. This analytical strategy is the most commonly used in studies both inside and outside of Spain. In addition to containing questions that are relevant for the research, the only requisite that these surveys must meet is that they must contain a sufficient number of cases to obtain a minimum of birth cohorts and educational transitions.

The CIS survey has a total sample size of 8,265 cases. However, not all the cases can be used for the analysis. The survey was carried out in 2006 and includes the resident population 16 years of age and older (born in 1990 or before). The cases have to be limited to those that are at least

<sup>5</sup> Ordinal discrete choice models.

25 years of age, or, in other words, those born in 1981 or before (if not, we run the risk that those of 16 to 24 years of age have not advanced further in their educations simply because they have not had time to do so). By applying this limit to the cases, the sample is reduced by 14% (some 1,150 cases); however, this makes the analysis more precise. Although the sample size is not very large, for the type of analysis carried out, it is sufficient to adequately minimise the number of birth cohorts and educational transitions considered. In their analysis, Barone and Ruggera (2018) excluded the countries that did not reach 2,000 cases. Although broader surveys exist that can be used to do this type of analysis, such as the Sociodemographic Survey, the Labour Force Survey and the Living Conditions Survey from Spain's INE, the CIS survey is the only one that incorporates variables measuring students' abilities and motivations, as well as parents' preferences regarding their children's educations; hence, we have chosen this survey for the analysis.

The following variables are used in the analysis (shown in Tables 3 and 4 in the Appendix): the dependent variable is education level reached, currently the most commonly used indicator in stratification studies. Four categories are considered: having lower than basic obligatory education (T0); having completed obligatory education (T1)<sup>6</sup>, having completed post-obligatory secondary education or vocational training (T2)<sup>7</sup>; having completed a university degree (T3).

<sup>6</sup> This refers to successfully completing Spain's ESO, EGB or the elementary Baccalaureate, depending on when the individual studied. FP1 [FP = Vocational Training] from the LGE system is excluded. The coding for the CIS groups Basic FP with FP1 from the LGE system, thus, although it makes analytical sense to separate them, it is not possible to do so.

<sup>7</sup> Includes all of FP, including what is not post-obligatory secondary education, such as FP1 from the LGE system and advanced level FP from the LOGSE.

This classification reflects the maximum level of education reached by the individual. It must be noted that this is a cumulative process; that is, T2 includes those that reached T1, in T3 are those that reached T1 plus those that reached T2 plus T3 (in fact, for example, T3 can increase without there being a greater percentage of persons that having T2, have decided to continue through T3, simply because T2 or T1 have increased). For the process to stop being cumulative, we consider the transitions conditioned by having achieved the prior educational level (which, in reality, constitute different variables because the number of cases vary): T2/T1 refers to achieving T2 when T1 has already been reached; T3/T2 refers to reaching T3 when T2 has already been reached. This way of conceiving educational achievement, as a sequential educational decision has predominated in studies on the evolution of IEO since the studies by Mare (1980, 1981).

Birth cohorts are used to measure the passage of time and as a vehicle for examining the evolution of IEO. We have constructed four: 1908-1949, 1950-1959, 1960-1969 and 1970-1981. Due to sample size we could not create more cohorts (ideally, there should be ten of 10 years) and it must be kept in mind that the first cohort covers a period of more than 40 years, while the other cohorts cover ten years. The use of birth cohorts is a standard method in these types of studies.

Regarding the main independent variable, social class of origin, we constructed a classification adapting the EGP scheme (Erickson-Goldthorpe-Protocarero)<sup>8</sup> to Spain and to the sample size (which does not permit many categories). This Weberian scheme of classes is constructed based on individuals' occupations and is the most used among sociologists studying social

<sup>8</sup> See Erickson, Goldthorpe and Portocarero (1979), as well as Erikson and Goldthorpe (1992).

stratification. Six social classes have been constructed using information about the father's occupation when the child was 16 years of age: Managers, professionals and technicians (I+II); non-manual workers (II-lab); business owners and self-employed (IVab); skilled manual workers (V+VI); unskilled manual workers (VIIa), farmers and agricultural workers (IVc+VIIb)<sup>9</sup>.

Other independent variables used are student ability and parents' preferences regarding education. Student's ability is measured through a subjective variable, grades received at 16 years of age (or before if the student abandoned his/her studies) and takes five values: the lowest if the student had very poor grades and the highest if he or she had very good grades. Parent's preferences are incorporated through a variable that measures parental attitude toward their child's education at 16 years of age (or before in the case of early abandonment of school) using three categories: negative if the parents encouraged their children to abandon their studies early to begin working; indifferent if it were not important to the parents if their children continued their educations; and positive, when the parents encouraged their children to study more and to continue with their educations. These variables tend not to be used in this type of analysis (generally because this type of information is not available).

We have used two control variables that are common in these types of studies: gender, in standard format, and the country in which the student lived longest until reaching 16 years of age (differentiating between those that lived in Spain and elsewhere).

We have used logistic regression as the method of parameter estimation, fol-

lowing the "educational transitions model" of Mare, which distinguishes changes in rates of school enrolment from the association between student's social origin and educational achievement, in line with the analysis in Shavit and Blossfeld (1993)<sup>10</sup>. As with other studies, the use of *logit* models is based on a prior choice regarding how to interpret the parameters; a *logit* model is used because the parameters are interpreted as *odds ratios*, consistent with the sociological definition of inequality in educational opportunities. The *logit* and *probit* models make different assumptions regarding error distribution, logistical for *logit* and normal for *probit*. In fact, both models present very similar results, almost equal in samples used in the social sciences (Fernández-Mellizo, 2015).

## RESULTS

The results of the analysis are presented in two sections; in the first, we analyse the evolution of IEO in Spain during the 20<sup>th</sup> century, which was a period of great expansion in schooling; in the second, we attempt to correct for selection bias in the models by introducing specific family and personal characteristics of the students.

### The evolution of inequality in educational opportunities in Spain

Before presenting the results for the evolution of IEO, it is necessary to consider the context of an expanding education system. There is much evidence of a widespread expansion of education in the 20<sup>th</sup> century (Haim and Shavit, 2013). One study of 33 countries (the majority in Europe) concluded that there was a sustained expansion of ed-

<sup>9</sup> We cannot distinguish among large and small business owners so we have placed them together (with the exception of the agricultural sector) in one category. We are grateful for the help of Miguel Caínzos in the construction of the variable social class of origin.

<sup>10</sup> This also permits us to control for changes in occupational structure over time, very important in Spain.

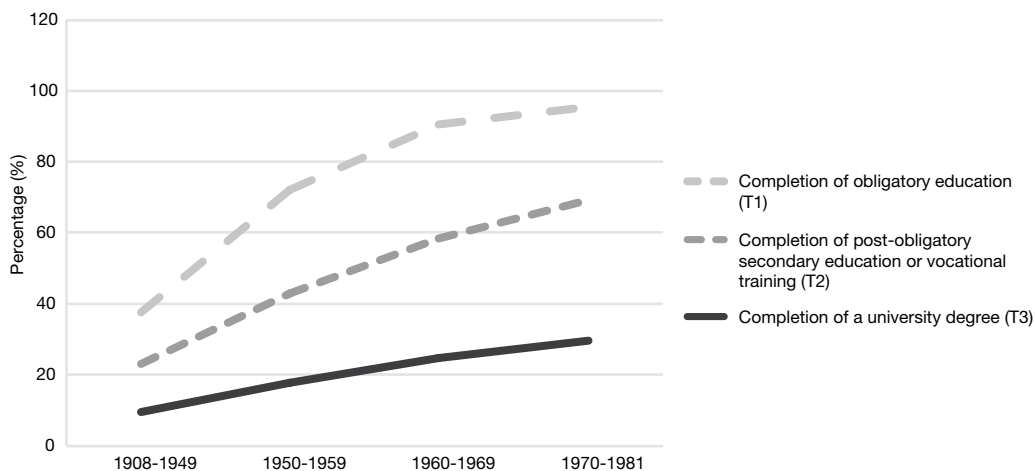
ucation throughout the period we are examining (Hadjar and Becker, 2016). The evidence in Spain for a continued expansion of education is also clear, lasting from the beginning of the 20<sup>th</sup> century until at least the mid-1970s (Fernández-Mellizo, 2014). The increase in the educational level of the Spanish population can be seen in historical statistics presented by the Fundación BBVA (Nuñez, 2005) and in other studies that use more recent data (Fuente and Doménech, 2016). Differences in findings are minor; some find that the expansion of education continued until the beginning of the 1980s (Paolo, 2012), while others describe a stagnation beginning at the end of the 1970s (Carabaña, 2013).

From a sociological perspective, the educational qualification reached is more important than the number of years of schooling, because the qualification is what permits labour market insertion and, ultimately, determines social position. As a re-

sult, in what follows we carry out an analysis of educational transitions, both of the highest level of education reached and of the conditioned transitions (upon having reached the prior educational level). As explained in the previous section, statistical analyses were carried out with four cohorts, as the survey is not sufficiently broad to analyse a greater number of cohorts. In comparing cohorts, it must be taken into account that the first cohort covers those born over a much longer period of time (four decades) than the other cohorts (which cover a period of more or less one decade).

Graph 1 shows the percentage of persons in each cohort that reached each level of education. Because of the sample size, as previously explained, we consider three educational levels: completion of obligatory education (T1), completion of post-obligatory secondary education or vocational training (T2) and completion of a university degree (T3).

**GRAPH 1.** *Percentage of persons in each cohort that reach a specific educational level*



Source: CIS study no 2634.

Looking at educational qualifications, we can see the expansion of education among those born in the 20<sup>th</sup> century in Spain. While only 38% of those born in

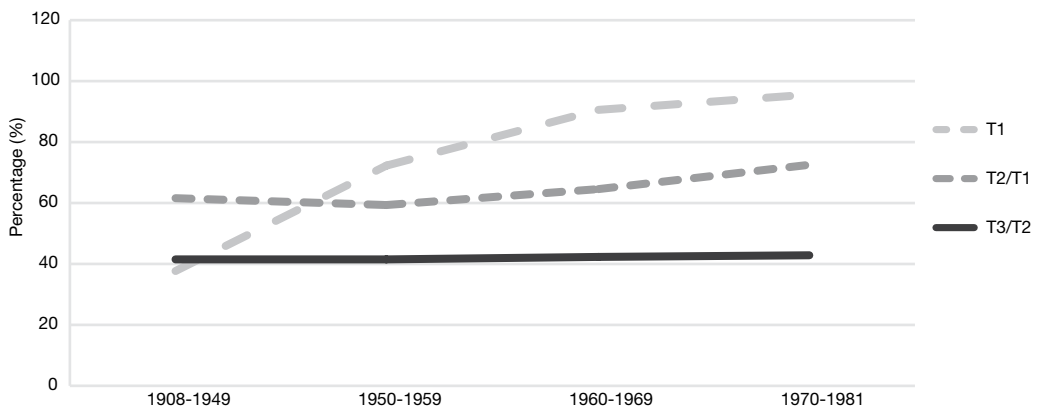
the first half of the century completed obligatory education, and only 23% post-obligatory secondary education or vocational training, and 10% received uni-

versity degrees, among those born after 1970, 96% completed obligatory education (almost reaching saturation at this education level), 69% post-obligatory secondary education (including VT) and 30% completed university studies. Thus, at all levels we find increases, although the graph also shows how the increases are smaller as the education level increases, that is, the expansion was more rapid at the two lower levels than at the highest one.

Graph 2 shows educational transitions conditioned by having obtained the prior level. Given that to pass to a specific education level it is necessary to have suc-

cessfully completed the prior level, it is interesting to know what percentage of those that have completed a specific education level reach the following level. This gives us a more accurate idea of the expansion of each education level separate from the effect of the expansion of the previous level. For example, to go through post-obligatory secondary education or vocational training, it is necessary to successfully complete basic obligatory education. It is of interest to know what percentage of those that complete the latter level continue (and successfully complete) post-obligatory secondary education or vocational training.

**GRAPH 2.** *Percentage that had the prior level of education that reaches the subsequent level (sequential decision) by cohort*



Notes: T1: Completion of obligatory education; T2: Completion of post-obligatory secondary education or vocational training; T3: Completion of a university degree; T2/T1: Complete T2 after T1 has been completed; T3/T2: Complete T3 after T2 has been completed.

Source: CIS study no 2634.

The percentage of those that complete obligatory education is the same as in the previous graph. However, the percentages of those that complete conditional educational transitions are higher than unconditional transitions, as those that do not pass to a higher level are left out of the calculations. For example, among those born in

1970 or after, 30% of the cohort successfully attained a university degree, while this percentage reaches 43% if we only take into account those that completed the previous education level, that is, those that successfully complete post-obligatory secondary education or vocational training. In other words, those that did not reach the

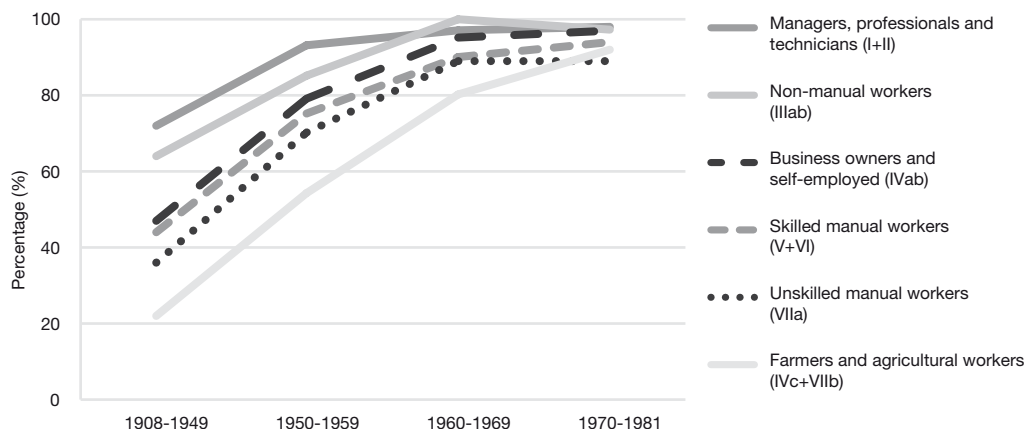
previous educational level, T2, are excluded from the calculation, given that it is impossible for them to study at the level of T3, which logically increases the percentage.

The increase across generations is very sharp for the case of obligatory education, as can be seen in Graph 2. However, the conditioned educational transitions (having reached the prior educational level) show a greater stability over time than in the case of transitions which do not establish conditions. The percentage of those that complete post-obligatory secondary education or vocational training (having completed the prior level) remains stable among those born in the first part of the 20<sup>th</sup> century, around 60%. In subsequent cohorts this percentage increases to 73% among those born in 1970 and after. More striking is the stability of the percentage of those that get university degrees (having completed the prior level), which remains at around 42% for the whole period considered.

To analyse the evolution of inequality in educational opportunities in Spain we intro-

duce social class of origin into the analysis based on the father's occupation when the student was 16 years of age. The analysis of the evolution of IEO shows the role that social origin plays in the models of educational transitions using non-linear techniques, such as logit, to obtain the estimated parameters. The impact of social class on obtaining educational qualifications is examined and, more specifically, in educational transitions to levels conditioned on having completed a prior level of education, as is standard practice in international research on this issue. To begin the analysis, we have carried out a descriptive analysis of the evolution by birth cohorts of the percentage of each social class that achieves each educational level. Graph 3 shows the percentage of each social class that successfully completes obligatory education, as well as the evolution of the latter through the different birth cohorts. We see a clear decline in differences across social classes, the result of a practical saturation of higher social classes, in the later birth cohorts.

**GRAPH 3.** *Percentage with obligatory education (T1) by social class of origin and cohort*



Source: CIS study no 2634.

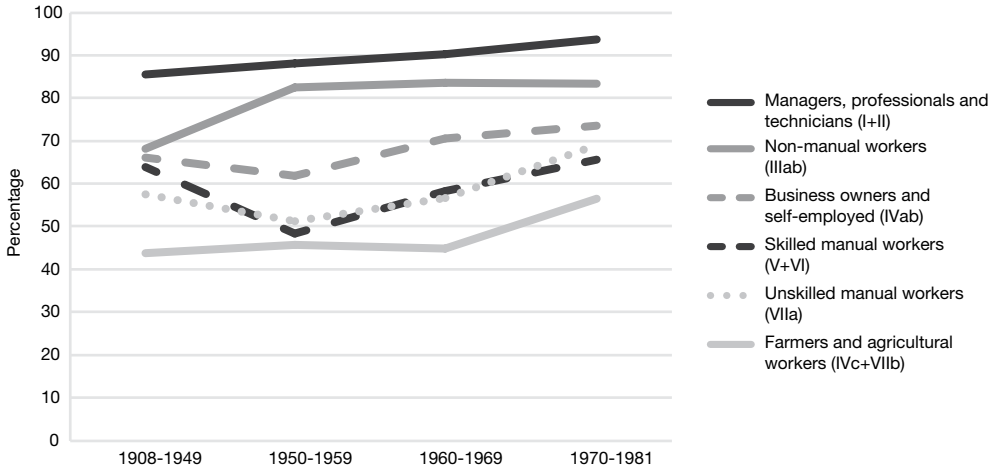
However, in the first of the conditional transitions, completing post-obligatory secondary education or vocational train-

ing after having completed obligatory education (T2/T1), we observe only a slight decrease in the differences among so-

cial classes, as can be seen in Graph 4. In the second of the conditional transitions, the percentage that obtains a university education after having successfully completed post-obligatory secondary educa-

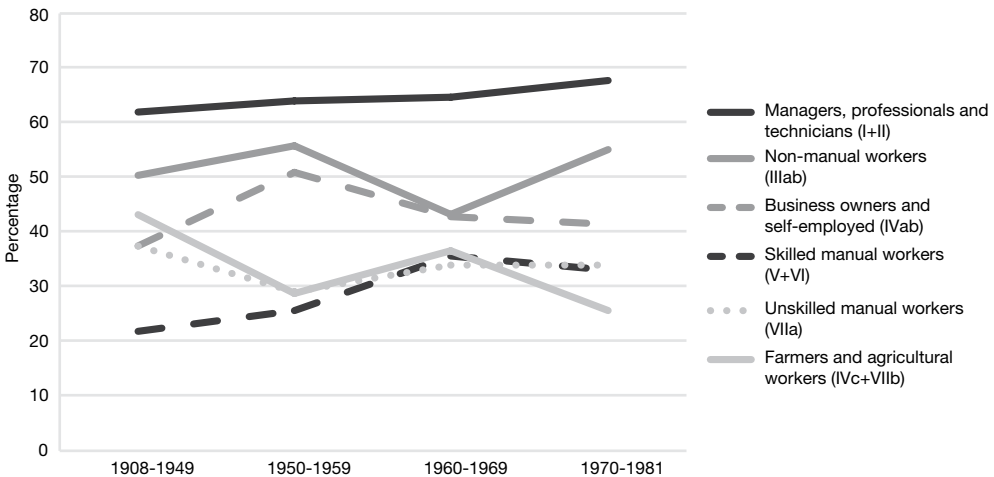
tion or vocational training (T3/T2), the evolution of the differences depends on each class in question and it is difficult to draw a general conclusion, as can be seen in Graph 5.

**GRAPH 4.** *Percentage with obligatory education that completed post-obligatory secondary education or VT (T2/T1) by social class and cohort*



Source: CIS study no 2634.

**GRAPH 5.** *Percentage that completed post-obligatory secondary education or VT with university education (T3/T2) by social class and cohort*



Source: CIS study no 2634.



In general, the graphs show no clear reduction in differences among social classes in terms of educational levels except in the case of the lowest education level (basic obligatory education, T1). The percentage completing obligatory education increased significantly during the period of time analysed and the higher social classes, as mentioned, are highly represented since the 1950s and, above all, 1960s. The decline in inequality that can be observed is specifically due to there being little room for increases among the higher classes.

To carry out an adequate statistical treatment of these data, we use logistical regression models, given that, as explained in the methodology section, the dependent variable is dichotomous. To simplify the analysis, following the international literature, we are only going to look at the conditional transitions (T2/T1 and T3/T2), in addition to T1. We include birth cohorts, the basis for comparison being the first time period, and social class, whose basis of comparison is the category of managers and professionals. We also include gender and the country in which individuals lived the greatest part of their lives up until 16 years of age as control variables. To see how IEO evolved for each social class, we include the interactions of the different cohorts with each social class. As shown in Table 1, in regard to T1, the expansion of education at this educational level during the whole period examined is clear, as seen in the positive (and significant) coefficients for the birth cohorts. The basis of comparison is the oldest cohort, so that a positive coefficient indicates that the following generations present higher educational achievements. Men and those that lived outside of Spain until 16 years of age also have greater likelihood of transitioning. The gender variable is measured as a dummy variable with a value of 1 if it is a man, so that the positive (significant) coefficient indicates greater probability of completing this transition in the case of men. The country variable is also a dummy with a value of 1 if the ma-

ajority of time until 16 years of age was spent living in Spain, so that a negative sign (significant) indicates that it is more likely that those that did not live in Spain made this transition. The impact of social class is also clear, as is shown by the (significant) negative coefficients, except for the children of non-manual workers, for whom the coefficient is not significant. The negative signs are logical given that the education level of each class is compared with that of the children of managers and professionals; if the level is lower the sign is, therefore, negative. Lastly, the stability of IEO in this transition is clear, as none of the coefficients for any of the interactions between birth cohort and social class are significant. Regarding T2/T1, the expansion of education is only found for the last cohort, that referring to those born in the 1970s, and gender is not significant. In addition, although the stability of IEO is dominant (interactions between the cohorts and social origin are not significant), we see an increase in inequality for skilled manual workers starting in the 1970s (the only significant coefficient). The latter is observed by a negative coefficient, which indicates that the differences in the educational achievements of the children of skilled manual workers in comparison to the children of managers and professionals declined as we advance from those born in the beginning of the 20<sup>th</sup> century until those born in the 1970s. Regarding T3/T2, there is no evidence of an expansion in education, and country of birth is also no longer important (no coefficients are significant). We again find stability in IEO during the whole period, except in the agricultural sector starting in the 1970s (where inequality increases). All the coefficients, except the latter, are not significant. Thus, the (significant) negative coefficient for children of agricultural workers born in the 1970s indicates that the difference in their educational achievement in comparison with the children of managers and professionals born at the beginning of the century increased.

**TABLE 1.** Logistic regression models to estimate reaching different educational levels

Variables	T1	T2/T1	T3/T2
<i>Cohorts (base: born 1908-1949)</i>			
Born 1950-1959	1.650***	0.080	0.100
Born 1960-1969	2.490***	0.320	0.140
Born 1970-1981	3.030***	0.810**	0.250
Gender	0.360***	0.070	0.000
Country	-0.480**	-1.030***	0.170
<i>Social class (base: managers and professionals)</i>			
Non-manual workers	-0.400	-1.110***	-0.420
Business owners and self-employed	-1.120***	-1.160***	-1.000***
Skilled manual workers	-1.220***	-1.260***	-1.770***
Unskilled manual workers	-1.530***	-1.590***	-1.060***
Agricultural sector	-2.290***	-2.060***	-0.770**
<i>Interaction year of birth and social class (base: born in 1908-1949; managers and professionals)</i>			
Born in 1950-1959; Non-manual workers	-0.490	0.760	0.050
Born in 1950-1959; Business owners and self-employed	-0.200	-0.270	0.450
Born in 1950-1959; Skilled manual workers	-0.340	-0.740	0.110
Born in 1950-1959; Unskilled manual workers	-0.270	-0.240	-0.420
Born in 1950-1959; Agricultural sector	-0.190	-0.070	-0.690
Born in 1960-1969; Non-manual workers		0.570	-0.480
Born in 1960-1969; Business owners and self-employed	0.480	-0.160	0.070
Born in 1960-1969; Skilled manual workers	-0.040	-0.580	0.540
Born in 1960-1969; Unskilled manual workers	0.150	-0.340	-0.260
Born in 1960-1969; Agricultural sector	0.190	-0.340	-0.390
Born in 1970-1981; Non-manual workers	-0.040	0.010	-0.100
Born in 1970-1981; Business owners and self-employed	0.440	-0.530	-0.070
Born in 1970-1981; Skilled manual workers	-0.030	-0.790*	0.310
Born in 1970-1981; Unskilled manual workers	0.530	-0.320	-0.350
Born in 1970-1981; Agricultural sector	0.607	-0.400	-1.020**
Constant	1.290***	2.790***	0.310
N	5,674	4,217	2,769
Pseudo-R <sup>2</sup>	0.320	0.080	0.050

Note: \* p<0.10; \*\* p<0.05; \*\*\* p<0.01.

Source: CIS study no 2634.

As can be seen in the above table, the expansion of education is clear in T1, but only is found for those born starting in the 1970s for T2/T3, while no expansion is found for T3/T2. In addition, characteristics such as gender and country in which students lived up until 16 years of age lose importance as education level increases. Regarding the influence of social class, its inclusion in the model

increases model fit in all the transitions (pseudo-R<sup>2</sup> increases)<sup>11</sup> and its impact is what was expected. In addition, we see how the impact of social class is different in function of education level (increases in passing from T1 to

<sup>11</sup> The tables in which the variables are included one by one are not shown here for reasons of space. They are available upon request.

T2/T1 and decreases in T3/T2), indicating that the main social selection is produced in T2). In relation to the evolution of inequality in educational opportunities, we find no evidence of any significant variation, given that the interactions with the birth cohorts are not significant (with the exception of an increase in inequality among those born starting in the 1970s among skilled manual workers for T2/T3 and in the agricultural sector for T3/T2). Therefore, the results of the analysis of educational transitions suggest stability in inequality of educational opportunities<sup>12</sup>.

### Trying to control for unobserved heterogeneity in the educational transitions models

Both the ability and motivation of students, and parents' preferences toward education, can have an impact on the likelihood of completing educational transitions. Not incorporating information on these variables, as is often the case in international literature and as is the case in the previous section, can bias the results of our analysis of the evolution of IEO by centring the analysis on a homogeneous group that is not representative of the overall population under study (those that continue studying). As Shavit and Blossfeld (1993) recognised, it would be best to have more information to be able to control for unobserved heterogeneity in the models. In fact, Mare (1993) uses a strategy similar to that used here (derived from information on siblings, not from direct questions as we have done here). Given that the CIS survey permits us to control for these variables, we are going

to replicate the analyses in the previous section with said controls, to see if the results are the same or not. Table 2 shows the results of this exercise.

In the first place, we see that both students' grades and the attitude of parents toward school are positively related with the probability of making each one of the educational transitions. The positive (significant) coefficients indicate that the higher the grades, the greater the academic achievement, and that if the parents have an indifferent or positive attitude toward education (in comparison with a negative attitude) the educational level reached is higher. The only exception is found in T3/T2 for the cases in which parents have an indifferent attitude toward schooling, in comparison with those that have a negative attitude (the coefficient, which is positive, is not significant). Secondly, the introduction of new information - students' grades and parents' preferences regarding education - improves the predictive capacity of all the models (pseudo r-squared increases), above all in T2/T1 and T3/T2. Lastly, the introduction of these new variables in the models reduces the impact of social class on educational transitions (the coefficients decrease)<sup>13</sup>. In the models without the mentioned controls, social class was picking up the effects of other omitted variables.

If we compare Table 2 with Table 1, in which we did not control for the new information introduced here, we see that the results of the analysis of the evolution of IEO are very similar; the great majority of interactions between social class and birth cohorts are not significant, providing additional findings favourable to the hypothesis of the stability of IEO. However, we do see some isolated significant interactions; in addition to those detected previously (increases in inequality in T2/T1 for skilled manual workers and in T3/T2 for the agricultural sector beginning with those

<sup>12</sup> Models were tested in which the cohort of reference was the second (those born in the decade of the 1950s), without detecting significant changes in the results. The reason for this simulation is that in the cohort by cohort analysis, not shown here, we see that the influence of social class in the probability of reaching T1 increases until we consider those born at the end of the 1950s and then declines until we reach those born at the beginning of the 1960s (see Table 5 in the Appendix).

<sup>13</sup> Detailed results are not shown here for reasons of space. They are available upon request.

born in the 1960s), we also see increases in T1 for non-manual workers born in the 1950s and in T2/T1 for skilled manual workers for any age cohort (negative and significant co-

efficients show, as previously explained, the lower achievements of this category in comparison to managers and professionals born in the first half of the 20<sup>th</sup> century).

**TABLE 2.** Logistic regression models to estimate educational achievement controlling for unobserved heterogeneity

Variables	T1	T2/T1	T3/T2
<i>Cohorts (base: born 1908-1949)</i>			
Born 1950-1959	2.070***	0.290	0.110
Born 1960-1969	2.690***	0.520	0.080
Born 1970-1981	2.890***	1.000**	0.220
Gender	0.470***	0.250***	0.110
Country	-0.300	-0.960***	0.280**
<i>Social class (base: managers and professionals)</i>			
Non-manual workers	0.080	-0.770*	-0.340
Business owners and self-employed	-0.080***	-0.880**	-0.920**
Skilled manual workers	-0.770***	-0.700*	-1.560***
Unskilled manual workers	-0.880***	-1.000**	-0.690
Agricultural sector	-1.710***	-1.600***	-0.670*
Grades	0.610***	0.880***	0.780***
<i>Parents' attitude toward education (base: negative)</i>			
Indifferent	0.340***	0.440**	0.110
Positive	1.200***	1.400***	1.310***
<i>Interaction year of birth and social class (base: born 1908-1949; managers and professionals)</i>			
Born in 1950-1959; Non-manual workers	-1.320*	0.460	0.020
Born in 1950-1959; Business owners and self-employed	-0.680	-0.320	0.390
Born in 1950-1959; Skilled manual workers	-0.810	-1.180**	-0.070
Born in 1950-1959; Unskilled manual workers	-0.750	-0.510	0.720
Born in 1950-1959; Agricultural sector	-0.600	-0.080	-0.640
Born in 1960-1969; Non-manual workers		0.600	-0.350
Born in 1960-1969; Business owners and self-employed	0.170	-0.380	-0.000
Born in 1960-1969; Skilled manual workers	-0.290	-0.910*	0.360
Born in 1960-1969; Unskilled manual workers	-0.090	-0.680	-0.500
Born in 1960-1969; Agricultural sector	-0.270	-0.530	-0.330
Born in 1970-1981; Non-manual workers	0.310	-0.010	0.000
Born in 1970-1981; Business owners and self-employed	0.550	-0.730	-0.120
Born in 1970-1981; Skilled manual workers	-0.110	-1.100**	0.160
Born in 1970-1981; Unskilled manual workers	-0.370	-0.650	-0.640
Born in 1970-1981; Agricultural sector	0.480	-0.640	-1.030**
Constant	-2.120***	-2.020	-4.120***
N	5,213	4,0770	2,712
Pseudo-R <sup>2</sup>	0.360	0.190	0.130

Note: \* p<0.10; \*\* p<0.05; \*\*\* p<0.01.

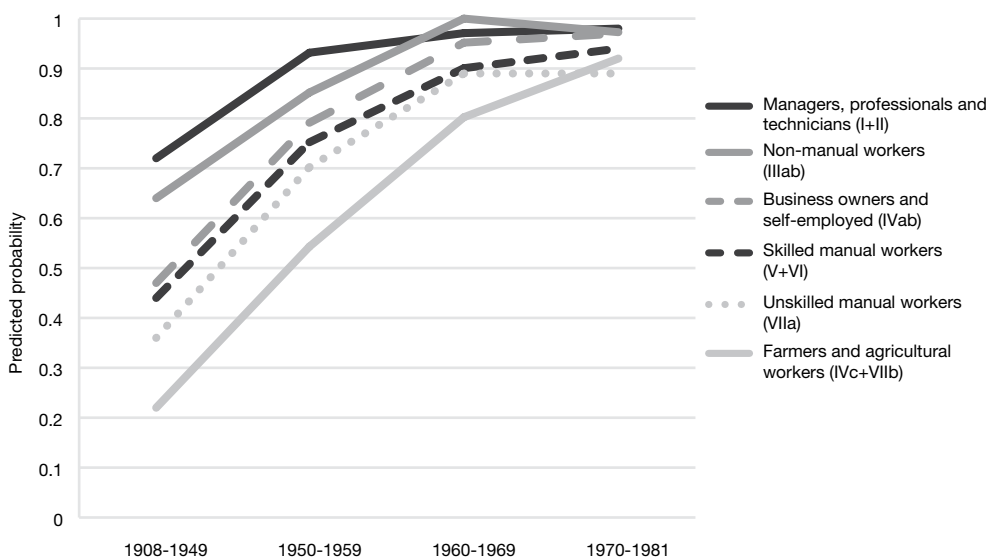
Source: CIS study no 2634.

We also see that in introducing the new information, gender is significant in T2/T1, being a man increases the likelihood of having completed secondary education, and country of birth in T3/T2, having been born in Spain increases the likelihood of having a university degree.

As in the previous section, we have replicated the analysis using the second cohort (those born in the 1950s) as the reference. In addition, we analyse ability (grades) and parents' educational preferences separately. The only difference in the analyses here is that in the models that only include the variable of students' grades and for which the

reference cohort is the second one, a significant interaction for T1 is found that shows a decline in inequality at this level for the agricultural sector among those born beginning in the decade of the 1970s (in comparison to those born in the 1950s), as indicated by the positive sign of the coefficient. The results of the analysis can be seen in Table 6 in the Appendix and in Graph 6, perhaps easier to interpret as it shows the estimate of the predicted marginal values derived from logistic models (at 95% confidence intervals). These values represent the predicted probability of completing obligatory education by social class and birth cohort.

**GRAPH 6.** *Probability of completing obligatory education according to social class by cohort (second cohort as basis of comparison)*



Source: CIS study 2634.

Therefore, we find that the results do not change substantially when introducing new information into the models; trends in inequality in educational opportunities remain fundamentally constant, as they did in the models without control. The only relevant change, a result of incorporating students' grades, is that we find a decrease in

inequality among farmers and agricultural workers (in relation to managers and professionals) born between the 1950s and the 1970s. The bias in the models of educational transitions that do not include said information, is not, therefore, as large as the literature consulted indicates (Shavit and Blossfeld, 1993; Mare, 1981; Mare, 1993;

Cameron and Heckman, 1998; Shavit, 2007; Bernardi, 2012). In any case, with the aim of being cautious, the possibility always exists that there is other important information that has not been considered.

## CONCLUSIONS

This article presents the results of a study on the evolution of inequality in educational opportunity in Spain, using data from a CIS survey that permits us to include younger cohorts and additional variables to obtain less biased results. We know that during the 20<sup>th</sup> century, as happened in other European countries and elsewhere, the overall educational level of the Spanish population increased. However, educational expansion was greatest at the lowest educational level (obligatory general education), lower at intermediate levels (post-obligatory secondary education and vocational training) and almost non-existent at the university level.

The central aspect of this study is the measurement of the evolution of IEO in Spain. It must be noted that the number of cases from the survey used is limited, and that the methodology employed also has its limitations. Nevertheless, the analysis carried out finds stability in IEO in the period of time considered, only finding certain isolated changes, in particular a rise in social class in specific cohorts (increases in inequality for those born in the 1970s in both T2/T1 for skilled manual workers and in T3/T2 for the agricultural sector). The expansion of education in Spain has not, therefore, led to a decline in inequality in education by social class. The average level of education of the population increased over the period studied, but inequalities by social origin continued throughout the period, as we suggested in our first hypothesis. This result is consistent with that found for Spain using data from the Sociodemographic Survey in a study by Martínez García (2002),

but not found by other authors; Ballarino et al (2009) found a decrease in IEO, while Peruga and Torres (1997) and Carabaña (1999) found both increases and decreases in inequality. Outside of Spain, the results of this study support the conclusions of Shavit and Blossfeld (1993) on the persistence of differences by social class in educational achievement over time and, in contrast, do not support the conclusions of Breen et al (2009) and Barone and Ruggera (2018), who found a decline in class differences. Thus, we find evidence that suggests that, in Spain, the process of educational expansion experienced by those born in the 20<sup>th</sup> century (up until the 1980s under democracy) was not accompanied by a decline in inequality in educational achievement by social origin. Rather, said process which lifted the general education level of the population coexisted with the persistence and maintenance of educational differences across social classes.

The most innovative part of the research is the control of unobserved heterogeneity in the models we used to analyse the evolution of IEO. In fact, this aspect is the basis for the use of data from the CIS survey, which permitted us to introduce variables measuring students' abilities and parents' preferences regarding their children's education. The incorporation of these new variables, absent in other studies, has allowed us to reduce possible selection bias (self-selection) found in other analyses of educational transitions (Shavit and Blossfeld, 1993; Mare, 1993), and obtain more trustworthy parameters. The results of the analysis confirm previous conclusions about the persistence of class differences in educational achievements, as variations in IEO were only found for isolated classes and cohorts (increases in inequality in T1 for non-manual workers born in the 1950s, increases in inequality in T2/T1 for skilled manual workers of any age cohort, and in T3/T2 for the agricultural sector born since the 1970s). However, although this

is the general conclusion, it must be noted that the inclusion of information on students' grades, reflecting their ability and motivation, modifies the models' predictions. In particular, at the lowest levels of education, incorporating students' grades reduces inequality in educational opportunities among students born between the 1950s and 1970s with parents in the agricultural sector. This supports the second hypothesis, though in a limited manner. By incorporating this information we are decreasing the selection bias that the educational system inevitably produces as only the better students remain as transitions happen; thus, in this way, we obtain less biased models. We cannot, however, discard the possibility that incorporating other factors could reduce existing bias even further. To the extent that new surveys incorporate more information we could develop better specified and more reliable models.

## BIBLIOGRAPHY

- Ballarino, Gabriele; Bernardi, Fabrizio; Requena, Miguel and Schadee, Hans (2009). "Persistent Inequalities? Expansion of Education and Class Inequality in Italy and Spain". *European Sociological Review*, 25(1): 123-138. doi: 10.1093/esr/jcn031
- Barone, Carlo and Ruggera, Lucia (2018). "Educational Equalization Stalled? Trends in Inequality of Educational Opportunity between 1930 and 1980 across 26 European Nations". *European Societies*, 20: 1-25. doi: 10.1080/14616696.2017.1290265
- Bernardi, Fabrizio (2012). "Unequal Transitions: Selection Bias and the Compensatory Effect of Social Background in Educational Careers". *Research in Social Stratification and Mobility*, 30: 159-174. doi: 10.1016/j.rssm.2011.05.005
- Blau, Peter M. and Duncan, Otis D. (1967). *The American Occupational Structure*. New York: John Wiley and Sons.
- Boudon, Raymond (1974). *Education, Opportunity, and Social Inequality: Changing Prospects in Western Society*. New York: Wiley.
- Breen, Richard (2010). "Educational Expansion and Social Mobility in the 20th Century". *Social Forces*, 89(2): 365-388. doi: 10.1353/sof.2010.0076
- Breen, Richard and Goldthorpe, John H. (1997). "Explaining Educational Differentials. Towards a Formal Rational Action Theory". *Rationality and Society*, 9(3): 275-305. doi: 10.1177/104346397009003002
- Breen, Richard and Jonsson, Jan O. (2005). "Inequality of Opportunity in Comparative Perspective: Recent Research on Educational Attainment and Social Mobility". *Annual Review of Sociology*, 31: 223-243. doi: 10.1146/annurev.soc.31.041304.122232
- Breen, Richard; Luijkx, Ruud; Müller, Walter and Polak, Reinhard (2009). "Nonpersistent Inequality in Educational Attainment: Evidence from Eight European Countries". *American Journal of Sociology*, 114 (5): 1475-1521. doi: 10.1086/595951
- Cameron, Stephen V. and Heckman, James J. (1998). "Life Cycle Schooling and Dynamic Selection Bias: Models and Evidence for Five Cohorts of American". *Journal of Political Economy*, 106: 262-334. doi: 10.1086/250010
- Carabaña, Julio (1999). *Dos Estudios sobre Movilidad Intergeneracional*. Madrid: Fundación Argentaria.
- Carabaña, Julio (2013). "Crecimiento del bachillerato e igualdad de los años ochenta". *Revista de la Asociación de Sociología de la Educación*, 6(1): 6-31. Available at: <https://ojs.uv.es/index.php/RASE/article/view/8595>, access January 15, 2021.
- Di Paolo, Antonio (2012). "Parental Education and Family Characteristics: Educational Opportunities across Cohorts in Italy and Spain". *Revista de Economía Aplicada*, 58 (XX): 119-146. Available at: <https://www.redalyc.org/pdf/969/96924442005.pdf>, access January 15, 2021.
- Erikson, Richard (2019). "How Does Education Depend on Social Origin?". In: Becker, R. (ed.). *Research Handbook on the Sociology of Education*. Massachusetts: E-Elgar.
- Erikson, Robert and Goldthorpe, John H. (1992). *The Constant Flux: A Study of Class Mobility in Industrial Societies*. Oxford: Oxford UP.
- Erikson, Robert y Jonsson, Jan O. (1996). *Can Education be Equalized? The Swedish Case in Comparative Perspective*. Boulder, Colorado: Westview.
- Erikson, Robert; Goldthorpe, John H. and Portocarero, Lucienne (1979). "Intergenerational Class Mobility in Three Western European Societies: England, France and Sweden". *The British Journal of Sociology*, 30: 415-441. doi: 10.2307/589632
- Fernández-Mellizo, María (2014). "The Evolution of Inequality of Educational Opportunities: A

- Systematic Review of Analyses of the Spanish Case”/“La evolución de la desigualdad de oportunidades educativas: una revisión sistemática de los análisis del caso español”. *Revista Española de Investigaciones Sociológicas*, 147: 107-120. doi: 10.5477/cis/reis.147.107
- Fernández-Mellizo, María (2015). “Continuidad o cambio en la desigualdad de oportunidades educativas: evidencia internacional y teorías”. *Revista Española de Sociología*, 23: 151-164. Available at: <https://recyt.fecyt.es/index.php/res/article/view/65378>
- Fuente, Ángel de la and Doménech, Rafael (2016). “El nivel educativo de la población en España y sus regiones: 1960-2011”. *Investigaciones Regionales*, 34: 73-94. Available at: <https://investigacionesregionales.org/es/article/el-nivel-educativo-de-la-poblacion-en-espana-y-sus-regiones-1960-2011/>
- Goldthorpe, John H. (2007). *On Sociology: Numbers, Narratives, and the Integration of Research and Theory*. Oxford: Oxford University Press. (2nd. edition).
- Hadjar, Andreas (2019). “Educational Expansion and Inequalities: How Did Inequalities by Social Origin and Gender Decrease in Modern Industrial Societies?”. In: Becker, R. (ed.). *Research Handbook on the Sociology of Education*. Massachusetts: E-Elgar.
- Hadjar, Andreas and Becker, Rolf (2016). “Education Systems and Meritocracy: Social Origin, Educational and Status Attainment”. In: Hadjar, A y Gross, C. (eds.). *Education systems and inequalities*. Bristol: Policy Press.
- Haim, Eyal B. and Shavit, Yossi (2013). “Expansion and Inequality of Educational Opportunity: A Comparative Study”. *Research in Social Stratification and Mobility*, 31: 22-31. doi: 10.1016/j.rssm.2012.10.001
- Mare, Robert D. (1980). “Social Background and School Continuation Decisions”. *Journal of the American Statistical Association*, 75(370): 295-305. doi: 10.1080/01621459.1980.10477466
- Mare, Robert D. (1981). “Change and Stability in Educational Stratification”. *American Sociological Review*, 46: 72-87. doi: 10.2307/2095027
- Mare, Robert D. (1993). “Educational Stratification on Observed and Unobserved Components of Family Background”. In: Shavit, Y. and Blossfeld, H. *Persistent Inequality. Changing Educational Attainment in Thirteen Countries*. Boulder: Westview.
- Martínez Celorrio, Xavier (2013). “Tendencias de movilidad y reproducción social por la educación en España”. *Revista de la Asociación de Sociología de la Educación*, 6(1): 32-41. Available at: <https://ojs.uv.es/index.php/RASE/article/view/8596>, access January 15, 2021.
- Martínez García, José S. (2002). *Habitus o calculus? Dos intentos de explicar la desigualdad de oportunidades educativas de los nacidos en España entre 1907 y 1966, con datos de la Encuesta Socio-Demográfica*. Carabaña, Julio (dir.), Madrid: Universidad Autónoma de Madrid. [Ph. Dissertation].
- Núñez, Clara Eugenia (2005). “Educación”. In: Carreras, A. y Tafunell, X. (coords.). *Estadísticas históricas de España. Siglos XIX-XX*. Bilbao: Fundación BBVA.
- Peruga, Rodrigo and Torres Mora, José A. (1997). “Desigualdad educativa en la España del siglo XX: un estudio empírico”. In: VV. AA. *Educación, vivienda e igualdad de oportunidades. II Simposio sobre Igualdad y Distribución de la Renta y la Riqueza*. Madrid: Fundación Argenteria-Visor.
- Raftery, Adrian E. and Hout, Michael (1993). “Maximally Maintained Inequality: Expansion, Reform, and Opportunity in Irish Education, 1921-75”. *Sociology of Education*, 66(1): 41-62. doi: 10.2307/2112784
- Shavit, Yossi and Blossfeld, Hans-Peter (1993). *Persistent Inequality. Changing Educational Attainment in Thirteen Countries*. Boulder: Westview.
- Shavit, Yossi; Arum, Richard and Gamoran, Adam (2007). *Stratification in Higher Education: A Comparative Study*. Palo Alto, California: Stanford University Press.
- Treiman, Donald J. (1970). “Industrialization and Social Stratification”. *Sociological Inquiry*, 40: 207-234. doi: 10.1111/j.1475-682X.1970.tb01009.x

**RECEPTION:** November 18, 2019

**REVIEW:** May 6, 2020

**ACCEPTANCE:** November 23, 2020



## APPENDIX

**TABLE 3.** *Descriptive quantitative and qualitative ordinal variables*

Variable	Type of variable	Observations	Average	Standard deviation	Minimum	Maximum
Number of years of schooling	Quantitative	7,338	11.58	4,59	0	25
Grades	Qualitative ordinal	7,306	3.58	0,84	1 (very poor)	5 (very good)
Year of birth	Quantitative	7,096	1,956.78	16,59	1,908	1,981

Source: CIS study 2634.

**TABLE 4.** *Descriptive qualitative nominal variables*

Variable	Type	Observations	Values	Frequencies
Gender	Dichotomous qualitative nominal	8,265	0: Woman 1: Man	2,917 3,106
Parents' attitude toward education	Qualitative nominal	7,222	1: Negative 2: Indifferent 3: Positive	827 1,158 5,236
Country of residence until 16 years of age	Dichotomous qualitative nominal	8,239	0: Other country 1: Spain	580 8,239
Social class of origin	Qualitative nominal	7,510	1: Managers and professionals 2: Non-manual workers 3: Business owners and self-employed 4: Skilled manual workers 5: Unskilled manual workers 6: Farmers and agricultural workers	806 518 1,272 1,754 1,095 2,061
Birth cohorts	Qualitative nominal	7,088	1: 1908-1949 2: 1950-1959 3: 1960-1969 4: 1970-1981	2,376 1,159 1,504 2,046
Highest level of education reached	Qualitative nominal	7,502	1: lower than basic obligatory education (T0) 2: Completed basic obligatory education (T1) 3: Completed post-obligatory secondary education or vocational training (T2) 4: Completed university degree (T3)	1,755 2,039 2,251 1,455
Reached T2 versus reaching only T1 (T2/T1)	Qualitative nominal dichotomous	5,751	0: Basic obligatory education (T1) 1: Post-obligatory secondary education or vocational training (T2)	2,040 3,710
Reached T3 versus reaching only T2 (T3/T2)	Qualitative nominal dichotomous	3,691	0: Post-obligatory secondary education or vocational training (T2) 1: University degree (T3)	2,241 1,449

Source: CIS study 2634.

**TABLE 5.** Logistic regression models to estimate T1 "cohort to cohort"

Variables	1908-1949	1950-1959	1960-1969	1970-1981
Gender	0.490***	0.550***	0.310*	-0.660***
Country	-2.180***	-0.920**	-1.190**	0.960***
<i>Social class (base: managers and professionals)</i>				
Non-manual workers	-0.380	-0.850		-0.440
Business owners and self-employed	-1.100***	-1.310***	-0.620	-0.650
Skilled manual workers	-1.210***	-1.540***	-1.240**	-1.340**
Un-skilled manual workers	-1.530***	-1.770***	-1.360**	-0.950
Farmers and agricultural workers	-2.270***	-2.480***	-2.080***	-1.580***
Constant	2.890***	3.270***	4.480***	3.680
N	1,616	984,000	1,209	1,795
Pseudo-R <sup>2</sup>	0.100	0.080	0.060	0.050

Note: \* p<0.10; \*\* p<0.05; \*\*\* p<0.01.

Source: CIS study 2634.

**TABLE 6.** Logistic regression models to estimate T1 controlling for unobserved heterogeneity (second cohort as base of comparison)

Variables	M1	M2	M3
<i>Cohorts (base: born 1950-1959)</i>			
Born 1908-1949	-1,400***	-1,450***	-2,060***
Born 1960-1969	1,270***	1,240***	0,670
Born 1970-1981	1,970***	2,000***	0,960
Gender	0,370***	0,460***	0,450***
Country	-0,510**	-0,280	-0,250
<i>Social class (base: managers and professionals)</i>			
Non-manual workers	-0,380	-0,180	-1,150*
Business owners and self-employed	-1,060***	-1,090***	-1,68***
Skilled manual workers	-1,330***	-1,270***	-1,840***
Un-skilled manual workers	-1,510***	-1,300***	-1,880***
Farmers and agricultural workers	-2,240***	-2,130***	-2,660***
Grades		0,700***	0,700***
<i>Interaction year of birth and social class (base: born 1950-1959)</i>			
Born in 1908-1949; Non-manual workers			1,030
Born in 1908-1949; business owners and self-employed			0,560
Born in 1908-1949; skilled manual workers			0,750
Born in 1908-1949; unskilled manual workers			0,570
Born in 1908-1949; farmers and agricultural workers			0,540
Born in 1960-1969; non-manual workers			
Born in 1960-1969; business owners and self-employed			0,840
Born in 1960-1969; skilled manual workers			0,520
Born in 1960-1969; unskilled manual workers			0,590
Born in 1960-1969; farmers and agricultural workers			0,420

**TABLE 6.** *Logistic regression models to estimate T1 controlling for unobserved heterogeneity (second cohort as base of comparison)* (Continuation)

Variables	M1	M2	M3
Born in 1970-1981; non-manual workers			1.140
Born in 1970-1981; business owners and self-employed			1.190
Born in 1970-1981; skilled manual workers			0.780
Born in 1970-1981; unskilled manual workers			1.240
Born in 1970-1981; farmers and agricultural workers			1.280*
Constant	2.710***	0.050	0.580
N	5,779	5,436	5,331
Pseudo-R <sup>2</sup>	0.320	0.340	0.330

Note: \* p<0.10; \*\* p<0.05; \*\*\* p<0.01.

Source: CIS study 2634.