



EMPIRIA. Revista de Metodología de las Ciencias Sociales

ISSN: 1139-5737

ISSN: 2174-0682

[empiria@poli.uned.es](mailto:empiria@poli.uned.es)

Universidad Nacional de Educación a Distancia

España

Valdés, Manuel Tomás

Principales procedimientos metodológicos para el análisis de la composición de la desigualdad educativa

EMPIRIA. Revista de Metodología de las Ciencias Sociales,  
núm. 48, 2020, Septiembre-Diciembre, pp. 115-145

Universidad Nacional de Educación a Distancia  
Madrid, España

DOI: <https://doi.org/10.5944/empiria.48.2020.28073>

Disponible en: <https://www.redalyc.org/articulo.oa?id=297169772005>

- Cómo citar el artículo
- Número completo
- Más información del artículo
- Página de la revista en [redalyc.org](https://www.redalyc.org)

[redalyc.org](https://www.redalyc.org)

Sistema de Información Científica Redalyc

Red de Revistas Científicas de América Latina y el Caribe, España y Portugal  
Proyecto académico sin fines de lucro, desarrollado bajo la iniciativa de acceso  
abierto

# *Principales procedimientos metodológicos para el análisis de la composición de la desigualdad educativa*

## *Main methodological procedures for the analysis of the composition of educational inequality*

MANUEL TOMÁS VALDÉS

Universidad Complutense de Madrid  
manueltv@ucm.es (ESPAÑA)

**Recibido:** 18.07 2018

**Aceptado:** 18.03.2020

### RESUMEN

El estudio de la desigualdad educativa en España se ha visto dificultado por la carencia de información longitudinal sobre transiciones educativas. Como resultado, distintos procedimientos metodológicos habituales en el estudio de las desigualdades educativas en el ámbito internacional han sido escasamente desarrollados en el caso español, lo que contribuye a un importante desconocimiento de los mismos. El presente artículo pretende ofrecer una revisión pedagógica de tales procedimientos, ejemplificando su uso con la expectativa de transición a la educación postobligatoria tal y como fue manifestada por el alumnado participante en las pruebas PISA en el año 2015.

En primer lugar, se presentan un conjunto de técnicas dirigidas a descomponer la desigualdad en una transición educativa (expectativa de transición en este caso) en un efecto primario (a través del rendimiento) y un efecto secundario (directamente sobre la toma de decisión). Aplicadas al caso de la expectativa de transición a la educación postobligatoria en España, tan solo el 40% de la desigualdad observada opera a través del rendimiento exhibido en las pruebas PISA. En segundo lugar, se introduce el método KHB, procedimiento dirigido a solucionar el problema del rescalamiento en modelos no lineales anidados y que, aplicado al estudio de la desigualdad educativa, permite poner a prueba la participación de mecanismos específicos de toma de decisión en la construcción de desigualdades. Tercero y último, se introduce el denominado modelo de compensación, donde los efectos secundarios del origen social no son constantes y se concentran en los niveles bajos de rendimiento. En efecto, se ha podido com-

probar que la mayor desigualdad en la expectativa de transición a la educación postobligatoria se observa en la parte baja de la distribución de rendimiento, siendo que dicho mecanismo de compensación da cuenta del 20% de la desigualdad observada.

## **PALABRAS CLAVE**

Desigualdad educativa, PISA, análisis contrafactual, método KHB, efectos primarios y secundarios, modelo de compensación.

## **ABSTRACT**

The study of educational inequality in Spain has been hindered by the lack of appropriate longitudinal datasets on educational transitions. As a result, different methodological procedures common in the study of educational inequality in the international context have been barely implemented in the Spanish case, which contributes to a widespread lack of awareness of their existence. This work intends to offer a pedagogical review of such procedures, exemplifying their use with the expectation of transition to post-compulsory education as it was manifested by the participants in 2015 PISA.

In the first place, I present a set of techniques aimed to decompose inequality in an educational transition (expectation of transition in this case) into a primary (via performance) and a secondary effect (directly on the decision-making). Applied to the expectation of transition to post-compulsory education in Spain, just 40% of inequality works through the performance exhibit in PISA. In the second place, I introduce the KHB method, aimed at addressing the rescaling problem in non-linear nested models which, applied to the study of educational inequality, makes it possible to test the participation of specific decision-making mechanisms in the construction of inequalities. Finally, I present the Compensatory Advantage model, where secondary effects are supposed to be non-constant and inequality is supposed to concentrate at low levels of performance. Indeed, we observe that the largest inequalities are found at the bottom of the performance distribution and that this compensation mechanism accounts for 20% of the observed inequality.

## **KEY WORDS**

Educational Inequality, PISA, counterfactual analysis, KHB method, primary and secondary effects, Compensatory Advantage model.

---

## 1. EL ESTUDIO DE LA DESIGUALDAD EDUCATIVA Y SU COMPOSICIÓN: EFECTOS PRIMARIOS Y SECUNDARIOS DEL ORIGEN SOCIAL

El medio escolar ha constituido desde hace décadas un prolífico campo de estudio para la Sociología dado el importante papel que el sistema educativo desempeña en los procesos de movilidad social. La transmisión intergeneracional de la posición en la estructura social puede ser corregida o alentada por la acción del sistema de enseñanza, y es en ese escenario donde se enmarca el análisis de la desigualdad educativa. Podemos definir su estudio como aquel que se ocupa de las diferencias en el logro educativo de los distintos grupos sociales, de donde se hace evidente que su magnitud y estabilidad temporal estará fuertemente condicionada por la forma en que conceptualicemos el logro educativo y modelicemos su análisis.

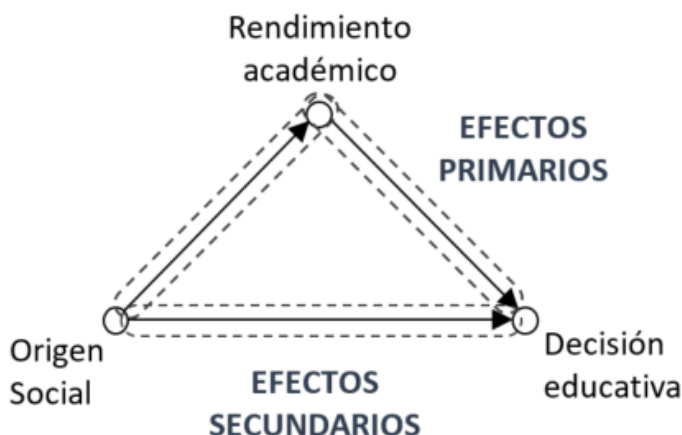
Durante décadas, el logro académico de un alumno se entendía apropiadamente resumido por los años de escolarización acumulados (Breen y Jonsson 2005). Piénsese que, en sistemas educativos poco complejos y previos a la expansión educativa, el número de años que el alumno hubiese permanecido en el sistema de enseñanza describía de forma bastante precisa el provecho que había sacado de su recorrido escolar. Conforme se intensificaron los procesos de expansión educativa y se diseñó una oferta formativa más diversificada en la educación secundaria superior y el nivel terciario, la validez del indicador fue puesta en cuestión (Martínez García 2011). A ello se añadían problemas metodológicos en el estudio de una variable continua como el número de años de escolarización acumulados, para cuyo análisis se requieren modelos de regresión lineal que confunden el efecto de la expansión educativa (el mayor espacio en términos de oportunidades que existe progresivamente dentro del sistema) con una auténtica reducción de la desigualdad educativa (la distinta propensión de las diferentes clases a aprovechar ese mayor espacio dentro del sistema). Es en dicho contexto en el que surge el denominado análisis de transiciones educativas, donde el recorrido escolar de un alumno es descrito como una sucesión de decisiones de continuidad o abandono analizadas a través de modelos logísticos que, por sus propias características, no confunden los efectos de la expansión educativa con la distinta propensión de las diferentes clases sociales a completar la transición bajo estudio (Mare 1980, 1981). En este nuevo escenario, la desigualdad educativa queda recogida por la distinta probabilidad de éxito en una particular transición educativa de alumnos de diferente extracción social, de la cual podremos estudiar su magnitud, su evolución temporal y, por lo que aquí interesa, su composición.

En el año 1974, el sociólogo francés Raymond Boudon popularizó la distinción entre los efectos primarios y secundarios del origen social. Por un lado, los alumnos de posiciones sociales más elevadas tienen una mayor probabilidad de completar cualquier transición educativa porque, en media, presentan un mayor rendimiento previo. El conjunto de recursos e influencias ejercidas en los hogares más acomodados impulsan el rendimiento de ese alumnado y, dado que

todos los alumnos se apoyan en su rendimiento previo para anticipar su probabilidad de éxito futuro, ese mayor rendimiento previo del alumnado de más alta extracción social se traduce en una mayor probabilidad de transición. A ello lo denominó Boudon efectos primarios del origen social, representados en la Figura 1 por el camino indirecto que relaciona el origen social (OS) con una decisión educativa (DE) a través del rendimiento académico (RA).

No obstante, tales efectos no dan cuenta del total de la desigualdad educativa. Ocurre que, ante situaciones de idéntico desempeño académico previo, alumnos pertenecientes a distintas clases sociales toman decisiones diferentes sobre su itinerario formativo; es decir, que, una vez controlado el rendimiento, la probabilidad de que un alumno de extracción social alta continúe estudios tras finalizar un nivel – o elija un itinerario más exigente – es superior a la que presentan los alumnos de extracción social baja. Boudon adjetivó tales efectos como secundarios, refiriendo la forma en que cada clase social se enfrenta y resuelve el proceso de toma de decisiones durante los momentos de transición educativa (Boudon 1974). Para ello, Boudon se apoyó en los trabajos de Keller y Zavalloni (1964), para quienes la ambición educativa es siempre relativa, dependiente de la distancia que separa la posición social del alumno de una cierta meta académica. De forma similar, Boudon (1998) argumentó que el principal factor que condiciona las decisiones educativas de los alumnos es su deseo de evitar la movilidad social descendiente, pero no todos los alumnos se ven igualmente expuestos a dicho riesgo ni requieren en la misma medida del sistema educativo para satisfacer ese objetivo: mientras que los alumnos de extracción social alta dependen en gran medida de su éxito académico para evitar el descenso social, los alumnos de extracción baja puede desarrollar itinerarios educativos más cortos y de menor exigencia sin incurrir en dicho riesgo.

**Figura 1.** Esquema de relaciones causales: efectos primarios y efectos secundarios.  
Fuente: Erikson, Goldthorpe, Jackson, Yaish y Cox (2005).



El estudio de la composición de la desigualdad ofrece resultados muy relevantes tanto desde el punto de vista investigador (permite comprender mejor los patrones de evolución de la desigualdad) como desde el punto de vista político (orienta las acciones dirigidas a reducir la desigualdad educativa) (Jackson 2013). No obstante, su aplicación en el caso español es anecdótica, lo cual no es de extrañar dada la vergonzante carencia de información longitudinal con que analizar transiciones educativas España.

## 2. OBJETIVOS Y FUENTE DE INFORMACIÓN

El interés que aquí se persigue es el de contribuir a la difusión de los principales procedimientos metodológicos con que abordar el estudio de la composición de la desigualdad educativa. En ese sentido, se pretende ofrecer un resumen pedagógico de tales procedimientos, evitando en la medida de lo posible un lenguaje excesivamente técnico y abundando en el detalle de lo realizado para permitir una adecuada comprensión.

No obstante, ya avanzaba que la precariedad de información en España es un hándicap mayúsculo. No es posible enfatizar suficiente la necesidad de mejorar la calidad de los datos disponibles para la investigación educativa en el caso español. Carecemos de una base de datos longitudinal que incluya información sobre el recorrido escolar del alumno (para analizar si completó o no la transición bajo estudio), su rendimiento académico en cada uno de los niveles matriculados (para conocer el rendimiento previo a dicha transición) y su origen social (incluyendo información sobre el nivel económico, social y cultural del hogar del encuestado durante el momento en que se enfrentó a la transición analizada). Contando con dicha información ha sido posible implementar los procedimientos que a continuación van a describirse en países como Alemania (Neugebauer y Schindler 2012; Zimmermann 2020), Países Bajos (Büchner y Van der Velden 2013; Kloosterman et al. 2009), Reino Unido (Jackson et al. 2007; Karlson 2013), Suecia (Erikson y Rudolph 2010), Rusia (Jackson, Khavenson, y Chirkina 2020), Italia (Barone, Triventi, y Assirelli 2018; Bernardi y Triventi 2020; Contini y Scagni 2013), Francia (Bernardi y Cebolla 2014a; Ichou y Vallet 2013), Estados Unidos (Morgan 2012; Morgan, Spiller, y Todd 2013) o Australia (Jerrim, Chmielewski, y Parker 2015), por citar tan solo algunos ejemplos.

En tanto que no contamos en España con tales bases de datos, ejemplificaré los procedimientos descritos empleando el estudio internacional PISA, que en distintas oleadas ha incorporado una pregunta sobre las expectativas educativas del alumnado. En efecto, y a pesar de que la mayoría de los trabajos citados han descompuesto la desigualdad observada en decisiones consumadas, no faltan ejemplos de estudios sobre la composición de la desigualdad en las expectativas y aspiraciones del alumnado (Neugebauer et al. 2013; Parker et al. 2016; Valdés 2019; Zimmermann 2020). En el presente trabajo emplearé la base de datos PISA 2015, donde la relación entre rendimiento, origen social y expectativas ya

ha sido analizada recientemente por Choi (2018), si bien no en los términos en que veremos más adelante.

Para aplicar los procedimientos que se van a describir son necesarias, al menos, las siguientes tres variables:

*Rendimiento.* PISA evalúa las competencias científica, lectora y matemática. A fin de llevar a cabo el análisis que nos atañe, he calculado la media de los resultados en las tres competencias y he obtenido los deciles de dicho indicador, agrupando a los alumnos según su pertenencia a cada uno de ellos.

*Origen social.* Para operacionalizar el origen social, he recurrido al Índice Socioeconómico y Cultural (ISEC) elaborado por la OCDE a partir del nivel formativo y estatus ocupacional de los progenitores y de distintas variables relativas a la posesión de bienes materiales y culturales en el hogar. A fin de ejemplificar los procedimientos de descomposición de forma sencilla, he construido una variable que distingue dos grupos sociales: uno de nivel socioeconómico y cultural bajo (primer quintil del ISEC, OS=0) y uno de nivel socioeconómico y cultural alto (quinto quintil del ISEC, OS=1).

*Expectativa de matriculación en educación postobligatoria (EPO).* A partir de la pregunta “¿Qué nivel de estudios esperas completar? Educación Secundaria Obligatoria (ESO), Ciclo Formativo de Grado Medio (CFGM), Bachillerato, Ciclo Formativo de Grado Superior (CFGS) o Universidad”, he construido una variable dicotómica con valor 0 cuando el alumno responde ESO (EPO=0) y 1 en caso contrario (EPO=1).

Antes de proceder con la exposición de los procedimientos, recuerde que el recurso a PISA es puramente ilustrativo, y no tiene mayor objetivo que ayudar a la mejor comprensión de las técnicas.

### **3. EFECTOS PRIMARIOS Y SECUNDARIOS DEL ORIGEN SOCIAL EN LA EXPECTATIVA DE MATRICULACIÓN EN EDUCACIÓN SECUNDARIA SUPERIOR EN ESPAÑA**

A partir de las tres variables anteriores podemos construir la Tabla 1, donde se desagrega la expectativa de transición a la EPO a partir del origen social del alumno y el decil de rendimiento en PISA.

---

**Tabla 1. Expectativa de matriculación en EPO por rendimiento en PISA y origen social del alumno.** Fuente: *Elaboración propia a partir de PISA 2015.*

		Decil de Rendimiento										Total
		1	2	3	4	5	6	7	8	9	10	
ISEC bajo	Total	1.586	1.282	991	937	727	647	517	419	365	208	7.679
	N	835	851	678	658	609	543	447	406	359	207	5.593
	EPO	%	53%	66%	68%	70%	84%	86%	97%	98%	99%	73%
ISEC alto	Total	180	283	354	469	631	727	983	1.030	1.336	1.740	7.733
	N	158	256	317	443	600	714	966	1.015	1.325	1.730	7.524
	EPO	%	88%	90%	90%	94%	95%	98%	98%	99%	99%	97%
Diferencia en la expectativa EPO		35%	24%	21%	24%	11%	14%	12%	2%	1%	0%	24%

De dicha tabla debe destacarse, en primer lugar, la existencia de una importante desigualdad en la expectativa de transición a la educación postobligatoria, en tanto que el 97% de los alumnos de ISEC alto espera matricularse en EPO, por el 73% de los alumnos de ISEC bajo. Una diferencia de 24 puntos porcentuales que pretendemos descomponer entre efectos primarios y secundarios.

En relación con los efectos primarios, podemos comprobar fácilmente que, mientras que el 53% del alumnado de ISEC alto se agrupa en los tres deciles de mejor rendimiento en PISA, el 50% del alumnado de ISEC bajo se concentra en los tres deciles de peor desempeño. Es decir, existe un efecto notable del origen social sobre el desarrollo competencial del alumnado español. Por otra parte, dentro de ambos grupos sociales es patente que, a mayor desarrollo competencial, mayor la probabilidad de formar la expectativa de matriculación en EPO. Así pues, parte de la desigualdad observada en dicha expectativa se deberá a la acción de efectos primarios.

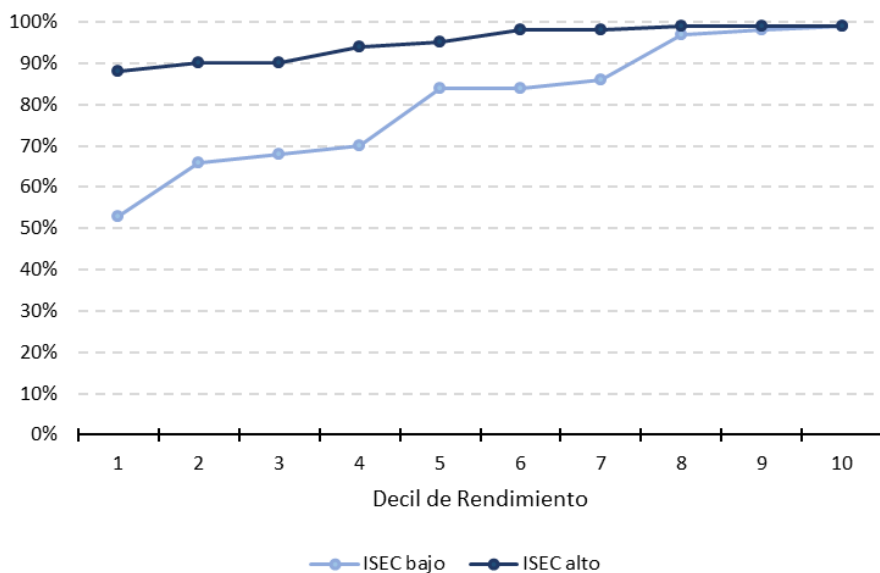
Dentro de cada decil, en cambio, los alumnos rinden de forma idéntica, lo que significa que la diferencia observada en la expectativa formativa entre ambos grupos sociales en cada decil se deberá únicamente a la distinta forma en que cada grupo social construye sus expectativas formativas, esto es, a los efectos secundarios. Como muestra la última fila de la Tabla 1, las diferencias entre grupos sociales dentro de cada decil son notables, especialmente en los siete primeros, alcanzando su máximo en el decil de rendimiento más bajo: 35 puntos porcentuales separan la expectativa de matriculación EPO del alumnado



de ISEC bajo (53%) y alto (88%) que ha alcanzado un idéntico desarrollo competencial (el correspondiente al decil 1).

El Gráfico 1 representa la expectativa de matriculación en EPO por origen social y rendimiento, donde la distancia vertical que separa ambas curvas en cada decil representa esos efectos secundarios. Es fácilmente apreciable la concentración de estos últimos en los niveles más bajos de desempeño, tal y como predice el denominado modelo de compensación (*Compensatory Advantage model*) que se introducirá después.

**Gráfico 1. Relación entre la expectativa de matriculación en EPO y el rendimiento en PISA. Fuente: Elaboración propia a partir de PISA 2015.**



Así pues, podemos concluir que la desigualdad en la expectativa de transición a la EPO del alumnado español se debe en parte a la acción de efectos primarios y en parte a la acción de efectos secundarios del origen social. Los siguientes procedimientos nos ayudarían a entender mejor la importancia relativa de ambos efectos, los mecanismos a través de los que se producen y la relación que existe entre ellos.

#### 4. PROCEDIMIENTOS DE ANÁLISIS DE LOS EFECTOS PRIMARIOS Y SECUNDARIOS DEL ORIGEN SOCIAL

He distinguido la exposición de los procedimientos en tres apartados. En el primero se presentan los procedimientos desarrollados durante la primera década del siglo XXI con el objetivo de obtener una medida de la importancia relativa de los efectos primarios y secundarios. Tales procedimientos están basados en el análisis contrafactual, donde la toma de decisión o la distribución de rendimiento de un grupo social es sustituida por la de otro grupo para ver cómo variaría la probabilidad de transición (expectativa de transición, en nuestro caso).

En el segundo apartado se describe el denominado método KHB, el cual permite comparar coeficientes en modelos anidados no lineales eludiendo el problema de reescalamiento. El método presenta una mayor sencillez que los procedimientos anteriores, pero, sobre todo, permite incorporar controles en la descomposición y testar mecanismos decisionales específicos en lugar de atribuir la desigualdad residual tras el control del rendimiento académico a los efectos secundarios.

Finalmente, en el tercer apartado introduciré el denominado modelo de compensación (*Compensatory Advantage model*), donde se deja atrás la asunción implícita en los procedimientos anteriores de que los efectos secundarios son constantes. El procedimiento descrito permite estimar el porcentaje de desigualdad que se debe específicamente al mecanismo de compensación.

##### 4.1. Escenarios factuales y contrafactuales en el cálculo de la importancia relativa de los efectos primarios y secundarios del origen social

Comencemos por el procedimiento planteado inicialmente por Erikson et al. (2005), y desarrollado posteriormente por Jackson et al. (2007) y Kartsonaki, Jackson y Cox (2013), donde se establecen escenarios contrafactuales en los que se conociese qué ocurriría si los alumnos de ISEC bajo, manteniendo su forma de tomar decisiones, hubiesen rendido como es característico entre los alumnos de ISEC alto; así como qué sucedería si esos mismos alumnos de ISEC bajo, rindiendo como suelen, tomaran decisiones como es propio entre los alumnos de ISEC alto. Sumados a los dos escenarios factuales en los que ambos grupos sociales rinden y deciden como les es característico, obtenemos los cuatro escenarios recogidos en la Tabla 2.

**Tabla 2. Escenarios factuales y contrafactuales**

	Decisión ISEC bajo	Decisión ISEC alto
Rendimiento ISEC bajo	A	B
Rendimiento ISEC alto	C	D

Obsérvese que, entre los escenarios recogidos en una misma columna (A-C y B-D), lo único que cambia es el rendimiento con el que se desempeñaron los alumnos, pues en ambos casos la toma de decisión permanece constante (no hay efectos secundarios). De la comparación de dichos escenarios podemos obtener una medida de efectos primarios. En cambio, entre los escenarios de una misma fila (A-B, y C-D) lo que varía es la toma de decisión, haciendo así posible conocer los efectos secundarios.

La cuestión es, por tanto, obtener la probabilidad de que un alumno supere la transición considerada en cada uno de esos escenarios. Dado que A y D ocurren en la realidad, es sencillo obtener tales probabilidades. Más problemático es calcular la probabilidad de transición en los escenarios contrafactuales C y B, para lo cual se han realizado distintas propuestas que veremos enseguida. Antes de ello, no obstante, completemos el procedimiento que llevaría finalmente a conocer la importancia relativa de los efectos primarios y secundarios del origen social.

Una vez obtenida la probabilidad de transición en cada escenario, calcularíamos los denominados *odds ratio*,  $Q_{jk}$ , donde el subíndice  $j$  indica el grupo social del que tomamos el rendimiento y el subíndice  $k$  el grupo social cuya forma de decidir consideremos. Por ejemplo,  $Q_{BA}$  representa el *odds ratio* de un hipotético alumno que rinde como lo hacen los alumnos de ISEC bajo y decide como lo hacen los alumnos de ISEC alto (escenario B).

$$Q_{jk} = \frac{P_{jk}}{1 - P_{jk}} \quad [1]$$

Para comparar escenarios entre sí se emplean los llamados *odds ratio* sintéticos (*synthesized odds ratio*). Su cálculo consiste sencillamente en el cociente de los *odds ratio* de dos escenarios. Así pues, la desigualdad observada en una transición educativa queda expresada por el *odds ratio* sintético que compara los escenarios A y D, es decir, el cociente entre  $Q_{BB}$  y  $Q_{AA}$ , que denotaremos como  $Q_{BB \cdot AA}$ , y que puede descomponerse de dos formas:

$$Q_{BB \cdot AA} = \frac{Q_{BB}}{Q_{AA}} = \frac{\frac{P_{BB}}{1 - P_{BB}}}{\frac{P_{AA}}{1 - P_{AA}}} = \frac{\frac{P_{BB}}{1 - P_{BB}}}{\frac{P_{BA}}{1 - P_{BA}}} \times \frac{\frac{P_{BA}}{1 - P_{BA}}}{\frac{P_{AA}}{1 - P_{AA}}} = Q_{BB \cdot BA} \times Q_{BA \cdot AA} \quad [2]$$

$$Q_{BB \cdot AA} = \frac{Q_{BB}}{Q_{AA}} = \frac{\frac{P_{BB}}{1 - P_{BB}}}{\frac{P_{AA}}{1 - P_{AA}}} = \frac{\frac{P_{BB}}{1 - P_{BB}}}{\frac{P_{AB}}{1 - P_{AB}}} \times \frac{\frac{P_{AB}}{1 - P_{AB}}}{\frac{P_{AA}}{1 - P_{AA}}} = Q_{BB \cdot AB} \times Q_{AB \cdot AA} \quad [3]$$

Empleando logaritmos,  $L_{jk} = \ln(Q_{jk})$ , podemos aprovechar sus propiedades para transformar estos productos en sumas, de manera que podamos expresar la desigualdad total (medida en *log odds*) como la acción agregable de un efecto primario y un efecto secundario.

$$L_{BB \cdot AA} = L_{BA \cdot AA} + L_{BB \cdot BA} \quad [4]$$

$$L_{BB \cdot AA} = L_{BB \cdot AB} + L_{AB \cdot AA} \quad [5]$$

En ambas expresiones, el primer sumando refleja los efectos primarios (compara escenarios incluidos en una misma columna), mientras que el segundo sumando recoge los efectos secundarios (compara escenarios incluidos en una misma fila). Para obtener la importancia relativa de ambos efectos solo será necesario, por tanto, dividir el correspondiente sumando entre el efecto total,  $L_{BB \cdot AA}$ . En tanto que las expresiones [4] y [5] pueden conducir a resultados distintos, es habitual el cálculo de su media aritmética (Erikson et al. 2005; Kartsonaki et al. 2013).

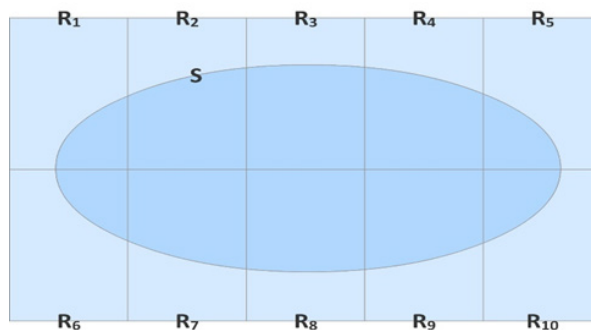
#### 4.1.1. Procedimiento de Erikson et al. (2005)

Como decíamos, para implementar dichos cálculos es necesario obtener la probabilidad de transición (expectativa de transición, en nuestro caso) en cada uno de los escenarios recogidos en la Tabla 2. Para ello, los propios Erikson et al. (2005) plantearon el siguiente modelo:

$$P_{jk}(EPO) = \int \left( \frac{1}{\sigma_j \sqrt{2\pi}} e^{-\frac{(x-\mu_j)^2}{2\sigma_j^2}} \right) \left( \frac{e^{a_k+b_k x}}{1 + e^{a_k+b_k x}} \right) dx \quad [6]$$

La expresión [6], por compleja que pueda parecer, no es más que una aplicación del Teorema de la Probabilidad Total, a partir del cual es posible calcular la probabilidad de un determinado suceso,  $S$ , a partir de la suma de sus intersecciones con un conjunto de sucesos,  $R_i$ , tal y cómo se muestra en la figura adjunta.

**Figura 2. Representación gráfica del Teorema de la Probabilidad Total**



Obsérvese que el espacio denotado como  $S$  es igual a la suma de la intersección de  $S$  con  $R_1$ , más la intersección de  $S$  con  $R_2$ , más la intersección de  $S$  con  $R_3$ , etc.

$$P(S) = P(S \cap R_1) + P(S \cap R_2) + \dots + P(S \cap R_{10}) = \sum_{i=1}^{10} P(S \cap R_i) \quad [7]$$

Sabemos además que la probabilidad de una intersección se corresponde con  $P(S \cap R_i) = P(R_i) \cdot P(S|R_i)$ , de modo que sustituyendo en [7] obtenemos que:

$$P(S) = \sum_{i=1}^{10} P(S \cap R_i) = \sum_{i=1}^{10} P(R_i) \cdot P(S|R_i) \quad [8]$$

Si, en vez de realizar un número finito de divisiones del espacio  $S$  (diez en la representación de la Figura 2), lo dividiésemos en fracciones infinitesimales, tendríamos que recurrir a una integral para agregarlas todas en lugar de a un sumatorio, de modo que:

$$P(S) = \int P(R_i) \cdot P(S|R_i) dR_i \quad [9]$$

Ahora piénsese sencillamente que el suceso  $S$  queda definido como “expectativa de transición a la EPO”, y que los distintos sucesos  $R_i$  se corresponden con el suceso “el alumno rinde al nivel  $R_i$  en PISA”. Diríamos entonces que  $P(S|R_i)$  es la probabilidad de que un alumno de rendimiento  $R_i$  espere realizar la transición a la EPO y  $P(R_i)$  la probabilidad de que los alumnos rindan al nivel  $R_i$  en PISA. Renombrando  $S$  como EPO y  $R_i$  como  $x$ , obtenemos que:

$$P(EPO) = \int P(x) \cdot P(EPO|x) dx \quad [10]$$

Comparando la anterior expresión [6] con esta expresión [10], podemos ver que Erikson et al. (2005) suponen que el rendimiento,  $x$ , sigue una distribución normal, y sustituyen  $P(x)$  por su función de distribución de probabilidad:

$$P(x) = \frac{1}{\sigma_j \sqrt{2\pi}} e^{-\frac{(x-\mu_j)^2}{2\sigma_j^2}} \quad [11]$$

Donde  $\mu_j$  y  $\sigma_j$  representan la media y la desviación típica del rendimiento del grupo social  $j$ . Esta asunción de normalidad es una de las debilidades principales de la propuesta pues, en ocasiones, la distribución del rendimiento puede presentar asimetrías que comprometan la precisión de los cálculos (Karlson y Holm 2011) o sencillamente no seguir una distribución normal (Hu 2017).

En cuanto al cálculo de la probabilidad de matriculación en EPO dado un determinado nivel de rendimiento previo,  $P(EPO|x)$ , Erikson et al. (2005) pro-

pusieron modelizar dicha probabilidad a través de una regresión logística, cuya ecuación se corresponde con:

$$P(ESS|x) = \frac{e^{a_k + b_k x}}{1 + e^{a_k + b_k x}} \quad [12]$$

Donde  $a_k$  y  $b_k$  representan la constante y el coeficiente del rendimiento previo en dicha regresión logística restringiendo el análisis para el grupo social  $k$ .

Obsérvese que cuando los subíndices  $j$  y  $k$  en la expresión [6] hacen referencia al mismo grupo social, el cálculo de la integral lleva a conocer los escenarios factuales (A y D), y cuando hacen referencia a grupos sociales distintos, permiten conocer los escenarios contrafactuales (B y C).

Volviendo al ejemplo de PISA 2015, los valores  $\mu$  y  $\sigma$  pueden obtenerse fácilmente calculando la media y la desviación típica del rendimiento de cada grupo social, a la vez que los valores  $a$  y  $b$  se obtienen ajustando el modelo de regresión logística para cada grupo social por separado. Tales parámetros se muestran en la Tabla 3 y resolviendo numéricamente la integral [6] a partir de cualquier software especializado obtenemos la probabilidad de expectativa de transición a la EPO en cada uno de los cuatro escenarios considerados.

Una manera sencilla de valorar la precisión del modelo es comparar la estimación que hemos realizado para los escenarios factuales con la proporción de alumnos de cada grupo social que espera matricularse en EPO. En el escenario A, donde el grupo de ISEC bajo rinde y decide como le es característico, ya dijimos que la probabilidad de matriculación en EPO era de 0,7284 (73% en la Tabla 1), habiéndose ahora obtenido un valor muy próximo de 0,7279. Y en el escenario D, donde el grupo de ISEC alto rinde y decide como suele, observamos una probabilidad de 0,9730 (97%) y ahora hemos calculado un valor de 0,9739. No parece, por tanto, que hayamos introducido importantes sesgos al asumir normalidad en la distribución del rendimiento ni al parametrizar la probabilidad de matriculación a través del modelo logístico.

**Tabla 3. Resultados de la aplicación del modelo de Erikson et al. (2005) para el cálculo de la probabilidad de expectativa de transición a la EPO en España.**

*Fuente: PISA 2015.*

Escenario	$\mu_j$	$\sigma_j$	$a_k$	$b_k$	P(EPO)
A	-0,5383	0,9527	1,5910	0,8680	0,7279
B	-0,5383	0,9527	3,3770	1,0070	0,9221
C	0,5897	0,8645	1,5910	0,8680	0,8706
D	0,5897	0,8645	3,3770	1,0070	0,9739

Aplicando los cálculos descritos en las expresiones [1] – [5], obtenemos que el 39,3% de la desigualdad observada en la expectativa de transición a la EPO es atribuible a efectos primarios (a las diferencias de rendimiento entre los alumnos de ISEC bajo y alto), mientras que el 60,7% restante se debe a efectos secundarios (a las desigualdades que se observan dentro de cada nivel de rendimiento).

A fin de evitarnos el tedioso trabajo de realizar estos cálculos manualmente, Kartsonaki (2010) ha desarrollado un paquete para R (DECIDE) que automatiza el proceso y, utilizando el método delta, permite comparar con los correspondientes intervalos de confianza los efectos primarios y secundarios en dos momentos del tiempo o en dos sistemas educativos, por ejemplo.

#### 4.1.2. Procedimiento de Buis (2010)

No obstante, ya avanzábamos que el anterior procedimiento se apoya en un supuesto de normalidad en la distribución del rendimiento que podría no ser cierto. A fin de soslayar tal dificultad, Buis (2010) repensó el procedimiento haciendo innecesario realizar asunciones sobre la forma en que se distribuye el desempeño de los alumnos.

En primer lugar, es necesario calcular un modelo de regresión logística para el conjunto de la muestra que explique la expectativa de transición a la EPO a partir del rendimiento académico, RA, el origen social, OS, y el efecto de interacción entre ambas variables independientes, RAxOS. Sustituyendo los coeficientes obtenidos usando la información de PISA 2015 en la ecuación del modelo de regresión logística, obtendríamos:

$$P(EPO = 1) = \frac{e^{1.591+0.868 \cdot (RA)+1.787 \cdot (OS)+0.138 \cdot (RA \times OS)}}{1 + e^{1.591+0.868 \cdot (RA)+1.787 \cdot (OS)+0.138 \cdot (RA \times OS)}} \quad [13]$$

Si sustituimos ahora los valores de rendimiento y origen social de cada alumno en la expresión [13], obtenemos su probabilidad estimada de matriculación en EPO. La media de tales probabilidades predichas entre los alumnos de ISEC bajo ofrece el escenario A, mientras que la media entre los alumnos de ISEC alto ofrece el escenario D.

El cálculo de los escenarios contrafactuales es algo más complejo. Pensemos primero en el escenario C, donde los alumnos de ISEC bajo rinden como es característico entre sus compañeros de ISEC alto. Dado que solo estamos interesados en ese alumnado de ISEC bajo, filtramos la muestra para considerar solo los casos en que OS = 0. Pero al sustituir los valores de cada alumno en la expresión [14], fijamos el valor OS en 1, asumiendo así que ese alumnado de ISEC bajo se comporta como sus compañeros de ISEC alto. La media aritmética de esas probabilidades estimadas ofrece la probabilidad de transición bajo el escenario C. Y de forma análoga procederíamos en el cálculo del escenario B, filtrando primero la muestra para considerar solo los casos en que OS = 1, y fijando después la variable OS en 0 para calcular las probabilidades estimadas de cada alumno.

Aplicando dicho procedimiento, se obtienen las probabilidades de expectativa de transición mostradas en la Tabla 4, y aplicando de nuevo los cálculos recogidos en las expresiones [1] – [5], se observa que el 38,8% de la desigualdad educativa observada es resultado de la acción de los efectos primarios, mientras que el 61,2% restante es imputable a la toma de decisión.

**Tabla 4. Escenarios factuales y contrafactuales siguiendo la propuesta de Buis (2010).** Fuente: PISA 2015.

	Decisión ISEC bajo	Decisión ISEC alto
Rendimiento ISEC bajo	0,7270	0,9228
Rendimiento ISEC alto	0,8699	0,9728

Buis (2010) ha elaborado un paquete para Stata (ldecomp) donde se automatiza el proceso descrito, el cual presenta la importante ventaja de admitir la incorporación de variables de control para la descomposición (género, raza, estatus migratorio,...).

#### 4.1.3. Procedimiento de Hu (2017)

Finalmente, Hu (2017) ha propuesto una nueva reformulación del método que no requiere realizar supuestos sobre la distribución del rendimiento, evita parametrizar la toma de decisión y permite además el trabajo con variables discretas de rendimiento. La propuesta es interesante no solo porque, en ocasiones, nos encontramos con bases de datos que no disponen de medidas continuas de desempeño, sino porque, además, la discretización del rendimiento es siempre posible y permite la aplicación de un modelo más sencillo de comprender y presentar que los propuestos por Erikson et al. (2005) y Buis (2010).

Disponiendo de esta variable de rendimiento categórica, Hu (2017) plantea que el deseo de matriculación en EPO puede ser expresado como la suma ponderada de la proporción de alumnos que en cada categoría de rendimiento esperan realizar dicha transición (proporción que refleja la toma de decisión y, por tanto, los efectos secundarios), utilizando como pesos la proporción que cada categoría de rendimiento representa sobre el total muestral (proporción que refleja el rendimiento y, por tanto, los efectos primarios):

$$P_{jk}(E+1) = \sum_{i=1}^R \left( \frac{n_{ij}}{N_j} \right) \left( \frac{m_{ik}}{n_{ik}} \right) \quad [14]$$

Donde el subíndice  $j$  hace referencia al grupo social del que tomemos el rendimiento y el subíndice  $k$  al grupo social cuya toma de decisión estemos considerando,  $R$  representa el número de grupos de rendimiento considerados,  $N_j$  el tamaño del grupo social  $j$ ,  $n_{ij}$  y  $n_{ik}$  el número de alumnos del grupo de rendimiento



$i$  en el grupo social  $j$  o  $k$ , respectivamente, y  $m_{ik}$  el número de alumnos del grupo de rendimiento  $i$  y el grupo social  $k$  que desean realizar la transición considerada.

Para comprobar a qué nos referimos con esa suma ponderada, recuperemos la expectativa de matriculación en EPO desagregada por decil de rendimiento y origen social que expusimos en la Tabla 1. Allí se dijo que el 73% de los alumnos de ISEC bajo y el 97% de los alumnos de ISEC alto esperaban matricularse en EPO calculando  $\frac{5.593}{7.679}$  y  $\frac{7.733}{7.524}$ , respectivamente. Pero a dichos resultados podríamos haber llegado a partir de las siguientes sumas ponderadas:

$$\frac{73}{100} = \frac{1.586}{7.679} \times \frac{835}{1.586} + \frac{1.282}{7.679} \times \frac{851}{1.282} + \frac{991}{7.679} \times \frac{678}{991} + \frac{937}{7.679} \times \frac{658}{937} + \frac{727}{7.679} \times \frac{609}{727} + \frac{647}{7.679} \times \frac{543}{647} + \frac{517}{7.679} \times \frac{447}{517} + \frac{419}{7.679} \times \frac{406}{419} + \frac{365}{7.679} \times \frac{359}{365} + \frac{208}{7.679} \times \frac{207}{208} \quad [15]$$

$$\frac{97}{100} = \frac{180}{7.733} \times \frac{158}{180} + \frac{283}{7.733} \times \frac{256}{283} + \frac{354}{7.733} \times \frac{317}{354} + \frac{469}{7.733} \times \frac{443}{469} + \frac{631}{7.733} \times \frac{600}{631} + \frac{727}{7.733} \times \frac{714}{727} + \frac{983}{7.733} \times \frac{966}{983} + \frac{1.030}{7.733} \times \frac{1.015}{1.030} + \frac{1.336}{7.733} \times \frac{1.325}{1.336} + \frac{1.740}{7.733} \times \frac{1.730}{1.740} \quad [16]$$

Donde las fracciones resaltadas en negrita representan la toma de decisión de cada grupo social (cuántos alumnos en cada decil de rendimiento quieren acceder a la EPO) y las fracciones no destacadas representan la distribución de su rendimiento (cuántos alumnos se encuentran en cada decil respecto del total de alumnos del grupo social considerado).

El lector observará que este proceder es análogo al descrito en el apartado 4.1.1, en tanto que puede ser igualmente derivado del Teorema de la Probabilidad Total. Recuérdesse la anterior expresión [8], donde supusimos que teníamos un número finito de grupos de rendimiento (diez deciles en este caso). Lo único que hace Hu es describir la probabilidad de que un alumno de rendimiento  $R_i$  y del grupo social  $k$  espere matricularse en EPO como:

$$P(EPO|R_i) = \frac{m_{ik}}{n_{ik}} \quad [17]$$

Mientras que la probabilidad de que un alumno del grupo social  $j$  pertenezca al decil de rendimiento  $R_i$  queda descrita por:

$$P(R_i) = \frac{n_{ij}}{N_j} \quad [18]$$

De esta forma, solo nos es necesario obtener las proporciones descritas en las expresiones [15] y [16], y combinarlas apropiadamente para calcular la probabilidad de matriculación en EPO en cada uno de los escenarios considerados (Tabla 5).

**Tabla 5. Resultados de la aplicación del modelo de Hu (2017) para el cálculo de la probabilidad de expectativa de transición a la EPO en España.**

		Decil de rendimiento										P(EPO)
		1	2	3	4	5	6	7	8	9	10	
A	$D_B$	0,53	0,66	0,68	0,70	0,84	0,84	0,86	0,97	0,98	1,00	72,84%
	$R_B$	0,21	0,17	0,13	0,12	0,09	0,08	0,07	0,05	0,05	0,03	
	$D_B * R_B$	0,11	0,11	0,09	0,09	0,08	0,07	0,06	0,05	0,05	0,03	
B	$D_A$	0,88	0,90	0,90	0,94	0,95	0,98	0,98	0,99	0,99	0,99	92,99%
	$R_A$	0,02	0,04	0,05	0,06	0,08	0,09	0,13	0,13	0,17	0,23	
	$D_A * R_A$	0,02	0,03	0,04	0,06	0,08	0,09	0,12	0,13	0,17	0,22	
C	$D_B$	0,53	0,66	0,68	0,70	0,84	0,84	0,86	0,97	0,98	1,00	89,05%
	$R_A$	0,02	0,04	0,05	0,06	0,08	0,09	0,13	0,13	0,17	0,23	
	$D_B * R_A$	0,01	0,02	0,03	0,04	0,07	0,08	0,11	0,13	0,17	0,22	
D	$D_A$	0,88	0,90	0,90	0,94	0,95	0,98	0,98	0,99	0,99	0,99	97,30%
	$R_B$	0,21	0,17	0,13	0,12	0,09	0,08	0,07	0,05	0,05	0,03	
	$D_A * R_B$	0,18	0,15	0,12	0,12	0,09	0,08	0,07	0,05	0,05	0,03	

Fuente: PISA 2015.

Aplicando los cálculos descritos en las expresiones [1] – [5], obtenemos que el 40,6% de la desigualdad educativa obedece al efecto del rendimiento previo, mientras que el restante 59,4% se explica por medio de la toma de decisión. Comparada con los dos anteriores procedimientos, la propuesta de Hu (2017) no es solo más inteligible, sino que presenta una mucho mayor facilidad de cálculo para el investigador menos familiarizado con la matemática y el análisis multivariable.

#### 4.2. Método KHB: efectos primarios y secundarios en modelos de regresión logística anidados

Uno de los principales problemas de los procedimientos anteriores es que se basan en una noción algo equívoca de efectos secundarios, en tanto que los asimilan a la desigualdad residual una vez controlado el rendimiento previo. Una manera más elegante de proceder sería operacionalizar mecanismos decisionales específicos e incorporarlos al modelo junto al rendimiento previo para observar qué parte de la desigualdad es atribuible a efectos primarios y qué parte a tales

mecanismos de decisión. Los anteriores procedimientos no permiten este tipo de análisis, pero sí sería posible desarrollarlo en un modelo de regresión al que progresivamente se le incorporan variables y se observa cómo se comporta el coeficiente del origen social.

Esta es una práctica sumamente extendida en investigación, donde se pretende valorar el efecto que una determinada variable independiente ( $x$ ) tiene sobre una variable dependiente ( $y$ ) tras tener en cuenta otras variables que denominamos mediadoras ( $z$ ). Para ello se calcula un primer modelo que solo incluye la variable independiente  $x$  (reducido) y un segundo modelo donde además añadimos  $z$  (completo). Se dice que los modelos están anidados porque el modelo completo es simplemente una ampliación del modelo reducido añadiendo la variable mediadora. De la comparación del coeficiente correspondiente a la variable  $x$  en ambos modelos ( $\beta_1^{M_I}$  y  $\beta_1^{M_C}$ ) deducimos tanto el efecto de control de  $z$  como el efecto de  $x$  sobre  $y$  una vez tenido en cuenta  $z$ .

$$M_C: y = \beta_0^{M_C} + \beta_1^{M_C} x + e \quad [19]$$

$$M_R: y = \beta_0^{M_R} + \beta_1^{M_R} x + \beta_2^{M_R} z + u \quad [20]$$

No obstante, dicha práctica, conocida comúnmente como análisis de mediación, no puede aplicarse directamente en modelos no lineales por el denominado problema de reescalamiento. Veamos en qué consiste. Denominemos  $Y^*$  a la propensión de un alumno de formar la expectativa de transición a la EPO, OS a su origen social y RA a su rendimiento académico. En tal caso, podemos construir los dos modelos siguientes (reducido y completo):

$$M_R: Y^* = \alpha_R + \beta_R SO \quad [21]$$

$$M_C: Y^* = \alpha_C + \beta_C SO + \gamma_C RA + \varepsilon \quad [22]$$

Dado que  $Y^*$  es una variable continua,  $M_C$  y  $M_R$  son modelos de regresión lineal, y comparar los coeficientes  $\beta_F$  y  $\beta_R$  nos ofrecería el efecto directo del origen social en la propensión de forma una expectativa educativa (efectos secundarios) y el efecto indirecto a través del rendimiento académico (efectos primarios). Sin embargo,  $Y^*$  no se puede observar porque los estudiantes nunca manifiestan su propensión a tomar una decisión o a formar una expectativa. Por tanto, los coeficientes en [21] y [22] no pueden ser estimados. En cambio, lo que sí observamos es la realización binaria de dicha propensión, en tanto que los alumnos finalmente manifiestan su expectativa en uno u otro sentido. De esta forma, debemos trabajar con una variable dicotómica en lugar de con una variable continua. Pese a ello, aún es posible construir los anteriores modelos anidados, aunque nos vemos obligados a hacer alguna asunción sobre la distribución y la varianza del término de error. Si asumimos una distribución logística binaria, entonces obtenemos que:

$$M_R: \text{logit}(Y) = a_R + b_R OS \quad [23]$$

$$M_C: \text{logit}(Y) = a_C + b_C OS + c_C RA + \varepsilon \quad [24]$$

Y puede demostrarse que:

$$b_C = \frac{\beta_C}{\sigma_C} \quad \text{y} \quad b_R = \frac{\beta_R}{\sigma_R}$$

Donde  $\sigma_C$  y  $\sigma_R$  son parámetros de escala que dependen de la varianza residual en los modelos lineales subyacentes reducido [21] y completo [22], respectivamente. En otras palabras, en esta especificación no lineal, cada modelo es escalado de acuerdo con la varianza residual de su modelo lineal subyacente. Cuando nosotros introducimos el rendimiento académico en la expresión [24], el coeficiente  $b$  cambia por dos razones. Por un lado, observamos un efecto de reescalamiento, dado que el rendimiento reduce la varianza residual del modelo subyacente. Por otro lado, observaremos un efecto de mediación, dado que parte del efecto del origen social sobre una decisión educativa opera a través del rendimiento académico. El problema en las especificaciones no lineales es que ambos efectos se confunden, y no es posible determinar a simple vista cuánto de la diferencia entre  $b_C$  and  $b_R$  se debe a cada uno de ellos (Allison 1999; Long y Mustillo 2018; Mize, Doan, y Long 2019; Mood 2010; Williams 2009).

Para abordar dicho problema y poder calcular los efectos primarios y secundarios del origen social en modelos anidados de regresión logística, Karlson, Holm y Breen desarrollaron el denominado método KHB (Breen, Karlson, y Holm 2018; Karlson 2013; Karlson y Holm 2011; Karlson, Holm, y Breen 2012; Kohler, Karlson, y Holm 2011). El método emplea los residuos (R) de la regresión del rendimiento académico (RA) sobre el origen social (OS):

$$RA = m + pOS + R \quad [25]$$

$$R = RA - (m + pOS) \quad [26]$$

Si incluimos R en la expresión [23], obtenemos el siguiente modelo reducido (reducido porque no se añade el rendimiento académico):

$$M_{\tilde{R}}: \text{logit}(Y) = \tilde{a}_R + \tilde{b}_R OS + \tilde{c}_R RA + \varepsilon \quad [27]$$

Dado que R absorbe toda la varianza del rendimiento académico no explicada por el origen social, la inclusión de R en  $M_{\tilde{R}}$  incrementa la varianza explicada hasta ser la misma que en  $M_C$ . Así pues, tenemos un nuevo modelo reducido,  $M_{\tilde{R}}$ , que, no obstante, está escalado por el mismo parámetro que  $M_C$  ( $\sigma_F = \sigma_{\tilde{R}}$ ). Así pues, los coeficientes en  $M_{\tilde{R}}$  y  $M_C$  son directamente comparables. En último término, la desigualdad total queda recogida por el coeficiente  $\tilde{b}_R$ , los efectos secundarios por el coeficiente  $b_F$  y los efectos primarios por la diferencia entre ambos,  $\tilde{b}_R - b_F$ .

Volviendo a nuestro ejemplo, la Tabla 6 muestra los tres modelos de regresión logística descritos en las expresiones [23], [24] y [27] para la expectativa de

transición a la EPO manifestada en PISA 2015. El primero incluye únicamente el origen social y nos permite conocer el efecto total que éste ejerce sobre la expectativa de matriculación en EPO. En la tercera columna se incluye además el control del rendimiento, permitiéndonos obtener una medida de efectos secundarios que, no obstante, estará afectada por el reescalamiento del modelo. Finalmente, en el modelo recogido en la segunda columna se incluye, en vez del rendimiento, la variable  $\widetilde{RA}$ , a partir de la cuál podremos conocer cuál es el efecto del origen social neto de reescalamiento. Obsérvese que la capacidad explicativa, y por tanto la varianza residual, de los modelos  $M_R$  y  $M_C$  es la misma, de modo que no hay reescalamiento entre ambos y podemos comparar sus respectivos coeficientes del origen social para obtener una medida del control que ejerce el rendimiento académico en la predicción de la expectativa de matriculación en EPO.

**Tabla 6. Modelos de regresión logística siguiendo la propuesta de Karlson, Holm, y Breen (2012) para abordar el problema del reescalamiento.**

Variable dependiente: expectativa de transición a la EPO			
	$M_R$	$M_{\widetilde{R}}$	$M_C$
Constante	0,987 (0,026)	1,131 (0,029)	1,611 (0,037)
Origen social, OS	2,593 (0,075)	2,753 (0,077)	1,748 (0,079)
Rendimiento académico, RA			0,891 (0,03)
Residuo del rendimiento, R		0,891 (0,03)	
R <sup>2</sup> Nagelkerke	22,00%	32,10%	32,10%

Fuente: PISA 2015.

De ignorar lo descrito hasta aquí, diríamos que el control del rendimiento provoca que el coeficiente  $b_1$  disminuya de 2,593 ( $M_R$ ) a 1,748 ( $M_C$ ), lo que significa unos efectos secundarios del 67,4% ( $\frac{1,748}{2,593}$ ). No obstante, esa medida se ve afectada por el reescalamiento, por lo que calculamos un nuevo modelo reducido ( $M_{\widetilde{R}}$ ) y repetimos el cálculo, observando ahora unos efectos secundarios que ascienden al 63,5% ( $\frac{1,646}{2,593}$ ), mientras que los efectos primarios se corresponderían con el 36,5% restante.

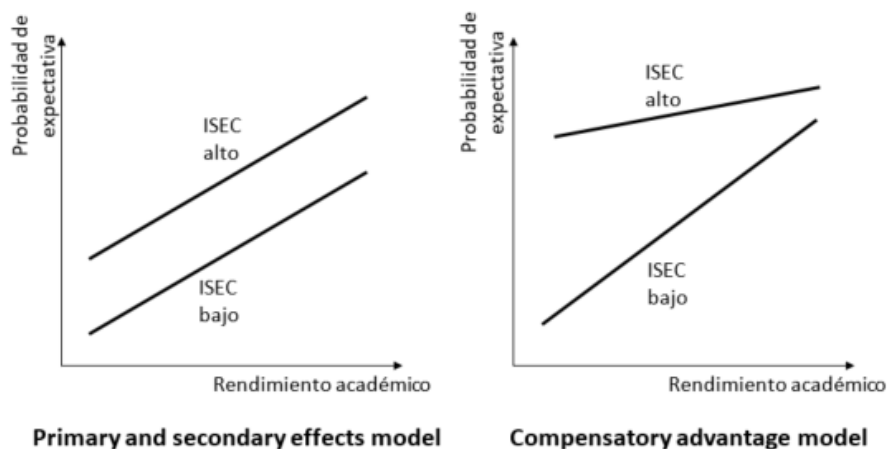
Por otra parte, el método KHB ofrece la ventaja de permitir la introducción de variables de control y, sobre todo, de mecanismos de toma de decisión específicos tales como la aversión al riesgo relativo (Breen y Goldthorpe 1997), las preferencias sobre el tiempo de retorno de la inversión en formación (Breen, van de Werfhorst, y Jæger 2014), el *habitus* del alumno (Bourdieu 1979), su

capital cultural (DiMaggio 1982) o la influencia de las expectativas y planes de *significant others* (Morgan 1998). De esta forma, en lugar de asimilar los efectos secundarios a la desigualdad residual tras el control del rendimiento, solo se les atribuiría el porcentaje del efecto del origen social que opera específicamente a través de tales mecanismos. Esta será sin duda una estimación menos generosa de los efectos secundarios, pero más precisa e informativa, en tanto que no solo nos indica el porcentaje de desigualdad directamente imputable a la toma de decisión, sino también los mecanismos específicos a través de los que esa desigualdad se produce, lo que es de sumo valor al diseñar programas de actuación para la reducción de las desigualdades ante la educación. El método KHB ha sido empleado en múltiples ocasiones con tales propósitos (Bachsleitner et al. 2018; Barone y Assirelli 2020; Barone et al. 2018; Fernández-Reino 2016; Jerrim et al. 2015; Lörz, Netz, y Quast 2016; Schindler 2017; Tjaden y Scharenberg 2017; Zimmermann 2020), para lo cual los autores del procedimiento, en colaboración con Kohler (2011), han desarrollado un paquete para Stata (*khb*) que facilita los anteriores cálculos.

### 4.3. Modelo de compensación: efectos secundarios heterogéneos

Finalmente, el denominado modelo de compensación (*Compensatory Advantage model*) pone en cuestión que los efectos primarios y secundarios sean aditivos o, lo que es lo mismo, que los efectos secundarios sean constantes a lo largo de la distribución del rendimiento. El modelo de compensación se centra en las estrategias desarrolladas por las familias de extracción social alta para lidiar con eventos adversos a lo largo de la vida que puedan afectar negativamente al logro escolar (Bernardi 2012, 2014). Aplicado al estudio de transiciones educativas, el mecanismo de compensación es conceptualmente equivalente a unos efectos secundarios heterogéneos, es decir, unos efectos secundarios que varían a lo largo de distribución de rendimiento y alcanzan su máximo entre el alumnado de peor desempeño (Bernardi y Cebolla 2014b, 2014a; Bernardi y Grätz 2015; Bernardi y Triventi 2020; Grätz y Wiborg 2020; Lehti, Erola, y Karhula 2019; Prix y Erola 2017; Troiano, Torrents, y Daza 2019).

La Figura 3 muestra la diferencia entre el modelo de efectos primarios y secundarios y el modelo de compensación. En este último, ambos grupos sociales responden de forma distinta al rendimiento cuando construyen sus expectativas formativas y toman sus decisiones educativas. Mientras que los alumnos de extracción social baja tienen muy en cuenta su rendimiento previo, los alumnos de extracción social alta toman decisiones independientemente de su desempeño anterior. Como resultado, la desigualdad se hace máxima en los niveles más bajos de rendimiento y tiende a desaparecer conforme mayor es el desempeño académico.

**Figura 3. Modelo de efectos primarios y modelo de compensación.**

Fuente: Bernardi y Cebolla (2014b).

Expresado de otra forma, el efecto del origen social sobre la formación de expectativas no sería constante, sino que variaría en función del rendimiento académico del alumno. Dicha situación es capturada en un modelo de regresión a partir de la interacción del origen social y rendimiento. Diremos que se observa un efecto de compensación cuando el coeficiente de dicha interacción sea negativo. La Tabla 7 muestra los resultados de un modelo de probabilidad lineal<sup>1</sup> para la expectativa de matriculación en EPO incorporando como predictores el origen social del alumno (ISEC del alumno estandarizado), su rendimiento (resultado en PISA estandarizado) y la interacción de ambas variables continuas. Como se puede observar, el efecto de interacción es negativo, y refleja el hecho, también ilustrado en el anterior Gráfico 1, de que los efectos secundarios se concentran en los niveles bajos de rendimiento y desaparecen a niveles altos.

<sup>1</sup> Como es sabido, los efectos de interacción son difícilmente interpretables a través de su coeficiente en especificaciones no lineales como un modelo logístico (Mize 2019), por lo que, en tales casos, es preferible recurrir a modelos de probabilidad lineal donde los coeficientes reflejan el efecto medio marginal asociado a cada variable (Mood 2010).

**Tabla 7. Modelo de probabilidad lineal para la expectativa de transición a la ESO incluyendo el efecto de interacción entre el origen social y el rendimiento académico.**

	Variable dependiente: expectativa de transición a la EPO		
	Coefficiente	Error estándar	Significatividad estadística
Constante	0.896	0.002	0.000
Origen social	0.042	0.002	0.000
Rendimiento académico	0.091	0.002	0.000
Origen social * Rendimiento académico	-0.036	0.001	0.000
R <sup>2</sup>	14,4%		

Fuente: PISA 2015.

Bernardi y Triventi (2020) han desarrollado recientemente un procedimiento que permite conocer cuánto de la desigualdad observada en una transición educativa (en nuestro caso, una expectativa de transición) es atribuible a este mecanismo de compensación. Para ello, es necesario distinguir al menos dos grupos sociales (recuperamos por tanto los alumnos de ISEC bajo y alto), y dividir la muestra en dos grupos de rendimiento (alumnos por debajo y por encima de la mediana de rendimiento en PISA).

Para empezar, la desigualdad en la expectativa de transición a la EPO no es más que la diferencia en la probabilidad de los alumnos de ISEC bajo y alto de formar dicha expectativa:

$$D^0 = P_B - P_A \quad [28]$$

Por otra parte, cada grupo social presenta una distribución característica de rendimiento. Por tanto, denominamos  $I_{AB}$  a la proporción de alumnos de ISEC alto con rendimiento bajo e  $I_{AA}$  a la proporción de alumnos de ISEC alto con rendimiento alto (nótese que  $I_{AB} + I_{AA} = 1$ ), de la misma forma que llamamos  $I_{BB}$  a la proporción de alumnos de ISEC bajo con rendimiento bajo e  $I_{BA}$  a la proporción de alumnos de ISEC bajo con rendimiento alto (de nuevo,  $I_{BB} + I_{BA} = 1$ ). La misma notación nos sirve para referirnos a la probabilidad de formar la expectativa de transición a la EPO en cada nivel de rendimiento y grupo social. Así pues, podemos expresar las anteriores probabilidades como:

$$P_B = (P_{AB} \times I_{AB}) + (P_{AA} \times I_{AA}) \quad [29]$$

$$P_A = (P_{BB} \times I_{BB}) + (P_{BA} \times I_{BA}) \quad [30]$$

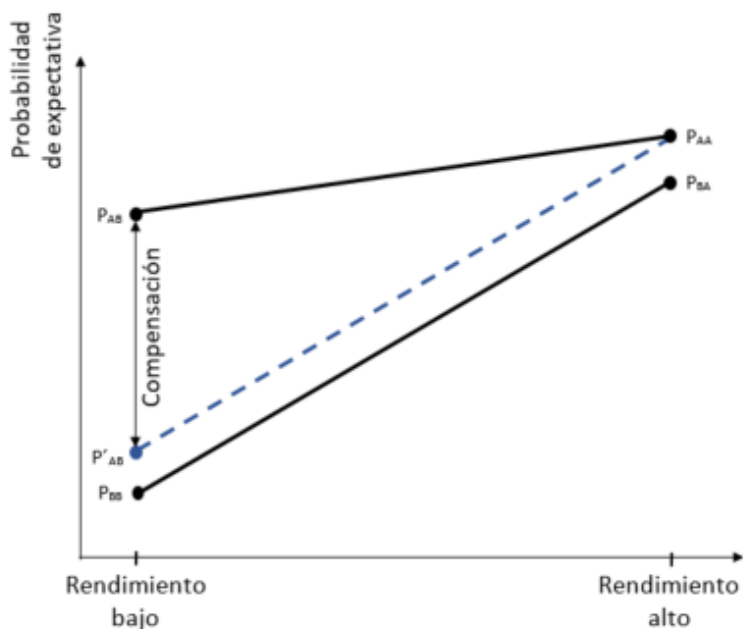
Y sustituyendo en [28], obtenemos que:

$$D^0 = [(P_{AB} \times I_{AB}) + (P_{AA} \times I_{AA})] - [(P_{BB} \times I_{BB}) + (P_{BA} \times I_{BA})] \quad [31]$$



El objetivo de Bernardi y Triventi (2020) es conocer cuánta desigualdad se hubiese observado de no haberse producido compensación, esto es, de no existir la observada interacción entre origen social y rendimiento académico. Como decíamos, las familias de ISEC alto compensan el bajo rendimiento mientras que las de ISEC bajo se resignan, pero en la parte más alta de la distribución de rendimiento no hay nada por lo que compensar. Por tanto, asumimos que en ese punto ( $P_{AA}$ ) no hay compensación, y proyectamos la pendiente de la curva del alumnado de ISEC bajo para observar qué probabilidad de formar la expectativa de transición a la EPO correspondería a los alumnos de ISEC alto en caso de no compensar el bajo rendimiento académico,  $P'_{AB}$ . La Figura 4, tomada directamente de Bernardi y Triventi (2020), ilustra la cuestión.

**Figura 4. Simulación de un escenario donde no se produce compensación del bajo rendimiento en la configuración de la expectativa de matriculación en la EPO**



Fuente: Bernardi y Triventi (2020).

Obsérvese que el punto  $P'_{AB}$  se corresponde con:

$$P'_{AB} = P_{BB} + (P_{AA} - P_{BA}) \quad [32]$$

De modo que podemos calcular la desigualdad generada en este escenario simulado sin compensación,  $D^S$ , sencillamente sustituyendo  $P_{AB}$  por esta nueva probabilidad  $P'_{AB}$ :

$$D^S = [(P_{LB} + (P_{UG} - P_{LG})) \times I_{UB}] + (P_{UG} \times I_{UG}) - [(P_{LB} \times I_{LB}) + (P_{LG} \times I_{LG})] \quad [33]$$

Finalmente, el porcentaje de desigualdad atribuible al mecanismo de compensación,  $\%D^C$ , es igual a la proporción que la diferencia  $D^0 - D^S$  representa sobre la desigualdad total,  $D^0$ :

$$\%D^C = \left[ \frac{(D^0 - D^S)}{D^0} \right] \times 100 \quad [34]$$

La Tabla 8 recoge tales cálculos aplicados al caso de la expectativa de matriculación en EPO. Para calcular las probabilidades empleadas en la expresión [33], utilizamos el modelo recogido en la Tabla 7 para estimar las probabilidades  $P_{AA}$ ,  $P_{AB}$ ,  $P_{BB}$  y  $P_{BA}$ . Las proporciones  $I_{AA}$ ,  $I_{AB}$ ,  $I_{BB}$  e  $I_{BA}$  se obtienen dividiendo el número de alumnos con rendimiento bajo y alto en cada grupo social entre el total de alumnos de cada grupo social. Recuérdese que, por tanto, debe cumplirse que  $I_{AB} + I_{AA} = 1$  y que  $I_{BB} + I_{BA} = 1$ . Sustituyendo tales valores en la expresión [33], obtenemos que, en un escenario donde las familias de ISEC alto no desarrollasen estrategias de compensación por el bajo rendimiento de sus hijos, la desigualdad hubiese sido de 19,8 puntos en lugar de los 24,4 puntos observados en realidad, es decir, que el mecanismo de compensación da cuenta del 19,2% de la desigualdad observada en la expectativa de matriculación en EPO en España.

**Tabla 8.** Simulación de un escenario en el que no se produce compensación por bajo rendimiento.

		P	I	$D^0$	$D^S$	$\%D^C$
ISEC bajo	Rendimiento bajo	0,66	0,72			
	Rendimiento alto	0,91	0,28			
				24,4%	19,8%	19,2%
ISEC alto	Rendimiento bajo	0,93	0,25			
	Rendimiento alto	0,99	0,75			

Fuente: PISA 2015.

## 5. CONCLUSIONES

En el presente trabajo hemos podido comprobar que el origen social condiciona notablemente la expectativa de transición a la Educación Postobligatoria en España donde 24 puntos porcentuales separan la expectativa de los alumnos de ISEC bajo y alto. Asimismo, hemos observado que esa desigualdad en la expectativa de transición puede descomponerse entre la acción indirecta del origen social sobre el rendimiento previo (efectos primarios) y la acción directa sobre la configuración de expectativas (efectos secundarios), siendo que tan solo

el 40% de la desigualdad en la expectativa de transición a la EPO en España se debe a la acción de efectos primarios. Lamentablemente, no disponemos en nuestro país de las bases de datos necesarias para trasladar tales análisis al estudio de transiciones educativas, lo que supone una limitación enorme que impide realizar trabajos homologables a los habituales en el contexto internacional.

Pese a ello, el estudio de la expectativa de transición a la EPO nos ha servido para describir distintos procedimientos dirigidos a analizar la composición de la desigualdad educativa. En primer lugar, hemos visto un grupo de técnicas que permiten conocer escenarios contrafactuales sobre la probabilidad de transición (o expectativa de transición, en nuestro caso). A partir de ello es posible obtener la importancia relativa de los efectos primarios y secundarios, lo que puede ser muy útil para analizar diversas hipótesis. Por ejemplo, podríamos estudiar la evolución en el tiempo de la composición de la desigualdad en términos de efectos primarios y secundarios, quizás como resultado de una reforma del sistema de becas o un cambio en el modelo de orientación escolar que deberían afectar a los efectos secundarios. O podríamos valorar hasta qué punto las diferencias en el coste de la matrícula universitaria dan cuenta del tamaño de los efectos secundarios en la transición a la universidad en las distintas regiones de nuestro país. En tales casos, los procedimientos presentados en el epígrafe 4.1 serían adecuados.

No obstante, no deben olvidarse las limitaciones propias del modelo de efectos primarios y secundarios. Por un lado, resulta poco elegante asimilar la desigualdad residual tras el control del rendimiento académico a los efectos secundarios. Ciertamente, la idea de efectos secundarios remite a mecanismos específicos que generan desigualdades en cada nivel de rendimiento. Por tanto, podemos preguntarnos hasta qué punto el alumnado emplea mecanismos concretos en distintas transiciones educativas. Quizás podría ocurrir que los alumnos de extracción social alta recurran en mayor medida a mecanismos propios de acción racional, mientras que los alumnos de extracción social baja se dejan llevar por su *habitus* y se comportan de forma más inercial. Para valorar la importancia relativa de esos mecanismos de decisión, el método KHB descrito en la sección 4.2 sería una opción adecuada.

Por otro lado, el modelo de efectos primarios y secundarios asume que ambos efectos son aditivos y no interactúan, mientras que nosotros hemos podido comprobar que este no es el caso para la expectativa de matriculación en la EPO en España. El modelo de compensación describe mejor el patrón observado, donde las familias de extracción social alta compensan niveles bajos de rendimiento académico y logran mantener expectativas ambiciosas, mientras que las familias de extracción social baja se resignan y aceptan itinerarios formativos más cortos y de menor exigencia académica. Aplicando el procedimiento desarrollado por Bernardi y Triventi (2020), hemos conocido que el mecanismo de compensación explica casi el 20% de la desigualdad en la expectativa de transición a la EPO. Lo que podríamos preguntarnos ahora es si esa compensación ha aumentado como resultado de la mayor participación del alumnado de extracción social baja en la educación postobligatoria.

---

En definitiva, el presente trabajo pretende contribuir al mejor conocimiento y dominio de distintos procedimientos metodológicos dirigidos al análisis de la composición de la desigualdad educativa, a la espera de que mejores bases de datos permitan emplearlos de forma extendida para poner a prueba muy distintas hipótesis al respecto de la toma de decisión en momentos de transición educativa.

## 6. BIBLIOGRAFÍA

- ALLISON, O. D. (1999): "Comparing logit and probit coefficients across groups", *Sociological Methods & Research*, 28(2), pp. 186-208. <https://doi.org/10.1177/0049124199028002003>
- BACHSLEITNER, A., BECKER, M., NEUMANN, M. y MAAZ, K. (2018): "Social Background Effects in the Transition to a Doctoral Degree – Empirical Evidence from a German Prospective Study", *Research in Social Stratification and Mobility*, 57, pp. 24-34. <https://doi.org/10.1016/j.rssm.2018.07.004>
- BARONE, C. y ASSIRELLI, G. (2020): "Gender Segregation in Higher Education: An Empirical Test of Seven Explanations", *Higher Education* 79(1), pp. 55-78. <https://doi.org/10.1007/s10734-019-00396-2>
- BARONE, C., TRIVENTI, M. y ASSIRELLI, G. (2018): "Explaining Social Inequalities in Access to University: A Test of Rational Choice Mechanisms in Italy", *European Sociological Review* 34(5), pp. <https://doi.org/554-69>. [10.1093/esr/jcy028](https://doi.org/10.1093/esr/jcy028)
- BERNARDI, F. (2012): "Unequal Transitions: Selection Bias and the Compensatory Effect of Social Background in Educational Careers", *Research in Social Stratification and Mobility* 30(2), pp. 159-74. <https://doi.org/10.1016/j.rssm.2011.05.005>
- BERNARDI, F. (2014): "Compensatory Advantage as a Mechanism of Educational Inequality: A Regression Discontinuity Based on Month of Birth", *Sociology of Education* 87(2), pp. 74-88. <https://doi.org/10.1177/0038040714524258>
- BERNARDI, F., y CEBOLLA, H. (2014a): "Previous School Results and Social Background: Compensation and Imperfect Information in Educational Transitions", *European Sociological Review*, 30(2), pp. 207-217. <https://doi.org/10.1093/esr/jct029>
- BERNARDI, F. y CEBOLLA, H. (2014b): "Social Class and School Performance as Predictors of Educational Paths in Spain", *Revista Española de Investigaciones Sociológicas*, 146, pp. 3-22. <https://doi.org/10.5477/cis/reis.146.3>
- BERNARDI, F. y GRÄTZ, M. 2015: "Making Up for an Unlucky Month of Birth in School: Causal Evidence on the Compensatory Advantage of Family Background in England", *Sociological Science*, 2, pp. 235-251. <https://doi.org/10.15195/v2.a12>
- BERNARDI, F. y TRIVENTI, M. 2020: "Compensatory Advantage in Educational Transitions: Trivial or Substantial? A Simulated Scenario Analysis", *Acta Sociologica*, 63(1), pp. 40-62. <https://doi.org/10.1177/0001699318780950>
- BOUDON, R. (1974): *Education, opportunity, and social inequality: Changing prospects in Western society*, New York, Wiley.
- BOUDON, R. (1998): "Social Mechanisms without Black Boxes", en *Social mechanisms: An analytical approach to social theory*, New York, Cambridge University Press, pp. 172-203.
- BOURDIEU, P. (1979): *La distinción. Criterio y bases sociales del gusto*, Madrid, Taurus.

- BREEN, R. y Goldthorpe, J. H. (1997): "Explaining Education Differentials: Towards a Formal Rational Action Theory", *Rationality and Society*, 9(3), pp. 275-305. <https://doi.org/10.1177/104346397009003002>
- BREEN, R., y JONSSON, J. O. (2005): "Inequality of Opportunity in Comparative Perspective: Recent Research on Educational Attainment and Social Mobility", *Annual Review of Sociology*, 31, pp. 223-243. <https://doi.org/10.1146/annurev.soc.31.041304.122232>
- BREEN, R., KARLSON, K. B. y HOLM, A. (2018). "A Note on a Reformulation of the KHB Method", *Sociological Methods and Research*. <https://doi.org/10.1177/0049124118789717>
- BREEN, R., VAN DE WERFHORST, H. G. y JÆGER, M. M. (2014): "Deciding under Doubt: A Theory of Risk Aversion, Time Discounting Preferences, and Educational Decision-Making", *European Sociological Review*, 30(2), pp.258-70. <https://doi.org/10.1093/esr/jcu039>
- BÜCHNER, C., y VAN DER VELDEN, R. (2013): "How social background affects educational attainment over time in the Netherlands", en *Determined to succeed? Performance vs choice in educational attainment*, California, Stanford University Press, pp. 89-115.
- BUIS, M. L. (2010): "Direct and indirect effects in a logit model", *The Stata journal*, 10(1), pp. 11-29. <https://doi.org/10.1177/1536867X1001000104>
- CHOI, A. (2018): "De padres a hijos: expectativas y rendimiento académico en España", *Presupuesto y Gasto Público*, 90, pp. 13-31.
- CONTINI, D., y SCAGNI, A. (2013): "Social-origin inequalities in educational careers in Italy", en *Determined to succeed? Performance vs choice in educational attainment*, California, Stanford University Press, pp. 149-184.
- DIMAGGIO, P. 1982. "Cultural Capital and School Success: The Impact of Status Culture Participation on the Grades of U.S. High School Students", *American Sociological Review*, 47(2), pp. 189. <https://doi.org/10.2307/2094962>
- ERIKSON, R., GOLDTHORPE, J. H., JACKSON, M., YAISH, M., & COX, D. R. (2005): "On class differentials in educational attainment", *Proceedings of the National Academy of Sciences*, 102(27), pp. 9730-9733. <https://doi.org/10.1073/pnas.0502433102>.
- ERIKSON, R., y RUDOLPHI, F. (2010): "Change in social selection to upper secondary school: Primary and secondary effects in Sweden", *European Sociological Review*, 26(3), pp. 291-305. <https://doi.org/10.1093/esr/jcp022>.
- FERNÁNDEZ-REINO, M. (2016): "Immigrant Optimism or Anticipated Discrimination? Explaining the First Educational Transition of Ethnic Minorities in England", *Research in Social Stratification and Mobility*, 46, pp. 141-56. <https://doi.org/10.1016/j.rssm.2016.08.007>
- GRÄTZ, M. y WIBORG, O. N. (2020): "Reinforcing at the Top or Compensating at the Bottom? Family Background and Academic Performance in Germany, Norway, and the United States", *European Sociological Review*. <https://doi.org/10.1093/esr/jcz069>
- HU, A. (2017): "Using a discretized measure of academic performance to approximate primary and secondary effects in inequality of educational opportunity", *Quality & Quantity*, 51(4), pp. 1627-1643. <https://doi.org/10.1007/s11135-016-0356-8>.
- ICHOU, M., y VALLET, L. A. (2013): "Academic achievement, tracking decisions, and their relative contribution to educational inequalities: Change over four decades in
-

- France”, en *Determined to succeed? Performance vs choice in educational attainment*, California, Stanford University Press, pp. 116-148.
- JACKSON, M. (2013): *Determined to Succeed? Performance vs Choice in Educational Attainment*, California, Stanford University Press.
- JACKSON, M., ERIKSON, R., GOLDTHORPE, J. H., y YAISH, M. (2007): “Primary and secondary effects in class differentials in educational attainment: The transition to A-Level courses in England and Wales”, *Acta Sociologica*, 50(3), pp. 211-229. <https://doi.org/10.1177/0001699307080926>.
- JACKSON, M., KHAVENSON, T. y CHIRKINA, T. (2020): “Raising the Stakes: Inequality and Testing in the Russian Education System”, *Social Forces*, 98(4), pp. 1613–1635. <https://doi.org/10.1093/sf/soz113>
- JERRIM, J., CHMIELEWSKI, A. K., y PARKER, P. (2015): “Socioeconomic inequality in access to high-status colleges: A cross-country comparison”, *Research in Social Stratification and Mobility*, 42, pp. 20-32. <https://doi.org/10.1016/j.rssm.2015.06.003>.
- KARLSON, K. B. (2013): “Summarizing primary and secondary effects”, *Research in Social Stratification and Mobility*, 33, pp. 72-82. <https://doi.org/10.1016/j.rssm.2013.01.001>.
- KARLSON, K. B., y HOLM, A. (2011): “Decomposing primary and secondary effects: A new decomposition method”, *Research in Social Stratification and Mobility*, 29(2), pp. 221-237. <https://doi.org/10.1016/j.rssm.2010.12.005>.
- KARLSON, K. B., HOLM, A., y BREEN, R. (2012): “Comparing Regression Coefficients Between Same-sample Nested Models Using Logit and Probit: A New Method”, *Sociological Methodology*, 42(1), pp. 286-313. <https://doi.org/10.1177/0081175012444861>.
- KARTSONAKI, C. (2010): “DECIDE: DEComposition of Indirect and Direct Effects”, R package.
- KARTSONAKI, C., JACKSON, M., y COX, D. R. (2013): “Primary and secondary effects. Some methodological issues”, en *Determined to succeed? Performance vs choice in educational attainment*, California, Stanford University Press, pp. 34-55.
- KELLER, S. y ZAVALLONI, M. (1964): “Ambition and Social Class: A Respecification”, *Social Forces*, 43(1), pp. 58-70. <https://doi.org/10.2307/2575967>
- KLOOSTERMAN, R., RUITER, S., DE GRAAF, P. M., y KRAAYKAMP, G. (2009): “Parental education, children’s performance and the transition to higher secondary education: trends in primary and secondary effects over five Dutch school cohorts (1965-99)”, *The British Journal of Sociology*, 60(2), pp. 377-398. <https://doi.org/10.1111/j.1468-4446.2009.01235.x>.
- KOHLER, U., KARLSON, K. B. y HOLM, A. (2011): “Comparing Coefficients of Nested Nonlinear Probability Models”, *The Stata Journal: Promoting Communications on Statistics and Stata*, 11(3), pp. 420-438. <https://doi.org/10.1177/1536867X1101100306>
- LEHTI, H., EROLA, J. y KARHULA, A. (2019): “The Heterogeneous Effects of Parental Unemployment on Siblings’ Educational Outcomes”, *Research in Social Stratification and Mobility*, 64. <https://doi.org/10.1016/j.rssm.2019.100439>
- LONG, J. S. y MUSTILLO, S. A. (2018): “Using Predictions and Marginal Effects to Compare Groups in Regression Models for Binary Outcomes”, *Sociological Methods & Research*. <https://doi.org/10.1177/0049124118799374>



- LÖRZ, M., NETZ, N. y QUAIST, H. (2016): "Why Do Students from Underprivileged Families Less Often Intend to Study Abroad?", *Higher Education*, 72(2), pp. 153-174. <https://doi.org/10.1007/s10734-015-9943-1>
- MARE, R. D. (1980): "Social Background and School Continuation Decisions", *Journal of the American Statistical Association*, 75(370), pp. 295-305. <https://doi.org/10.2307/2287448>.
- MARE, R. D. (1981): "Change and stability in educational stratification", *American Sociological Review*, 46(1), pp. 72-87.
- MARTÍNEZ GARCÍA, J. S. (2011): "Género y origen social: diferencias grandes en fracaso escolar administrativo y bajas en rendimiento educativo", *Revista de la Asociación de Sociología de la Educación*, 4(3), pp. 270-285.
- MIZE, T. D. (2019): "Best Practices for Estimating, Interpreting, and Presenting Nonlinear Interaction Effects", *Sociological Science*, 6, pp. 81-117. <https://doi.org/10.15195/V6.A4>
- MIZE, T. D., DOAN, L. y LONG, J. S. (2019): "A General Framework for Comparing Predictions and Marginal Effects across Models", *Sociological Methodology*, 49(1), pp. 152-189. <https://doi.org/10.1177/0081175019852763>
- MOOD, C. (2010): "Logistic Regression: Why We Cannot Do What We Think We Can Do, and What We Can Do About It", *European Sociological Review*, 26(1), pp. 67-82. <https://doi.org/10.1093/esr/jcp006>
- MORGAN, S. L. (1998): "Adolescent Educational Expectations: Rationalized, Fantastized, or Both?", *Rationality and Society*, 10(2), pp. 131-162. <https://doi.org/10.1177/104346398010002001>
- MORGAN, S. L. (2012): "Models of College Entry in the United States and the Challenges of Estimating Primary and Secondary Effects", *Sociological Methods & Research*, 41(1), pp. 17-56. <https://doi.org/10.1177/0049124112440797>.
- MORGAN, S. L., SPILLER, M. W., y TODD, J. J. (2013): "Class origins, high school graduation, and college entry in the United States", en *Determined to succeed? Performance vs choice in educational attainment*, California, Stanford University Press, pp. 279-305.
- NEUGEBAUER, M., REIMER, D., SCHINDLER, S. y STOCKÉ, V. (2013): "Inequality in Transitions to Secondary School and Tertiary Education in Germany", en *Determined to succeed? Performance vs choice in educational attainment*, California, Stanford University Press, pp. 89-115.
- NEUGEBAUER, M., y SCHINDLER, S. (2012): "Early transitions and tertiary enrolment: The cumulative impact of primary and secondary effects on entering university in Germany", *Acta Sociologica*, 55(1), pp. 19-36. <https://doi.org/10.1177/0001699311427747>.
- PARKER, P. D., JERRIM, J., SCHOON, I. y MARSH, H. W. (2016): "A Multination Study of Socioeconomic Inequality in Expectations for Progression to Higher Education: The Role of Between-School Tracking and Ability Stratification", *American Educational Research Journal*, 53(1), pp. 6-32. <https://doi.org/10.3102/0002831215621786>
- PRIX, I. y EROLA, J. (2017): "Does Death Really Make Us Equal? Educational Attainment and Resource Compensation after Paternal Death in Finland". *Social Science Research*, 64, pp. 171-183. <https://doi.org/10.1016/j.ssresearch.2016.10.012>
- SCHINDLER, S. (2017): "School Tracking, Educational Mobility and Inequality in German Secondary Education: Developments across Cohorts", *European Societies*, 19(1), pp. 28-48. <https://doi.org/10.1080/14616696.2016.1226373>
-

- TJADEN, J. D. y SCHARENBERG, K. (2017): "Ethnic Choice Effects at the Transition into Upper-Secondary Education in Switzerland". *Acta Sociologica*, 60(4), pp. 309-324. <https://doi.org/10.1177/0001699316679491>
- TROIANO, H., TORRENTS, D. y DAZA, L. (2019): "Compensation for Poor Performance through Social Background in Tertiary Education Choices", *Studies in Higher Education*. <https://doi.org/10.1080/03075079.2019.1666262>
- VALDÉS, M. T. (2019): "Efectos primarios y secundarios en la expectativa de matriculación universitaria: la desigualdad como reto del siglo XXI", *Revista Prisma Social* 25, pp. 332-358.
- WILLIAMS, R. (2009). "Using Heterogeneous Choice Models to Compare Logit and Probit Coefficients Across Groups", *Sociological Methods & Research*, 37(4), pp. 531-559. <https://doi.org/10.1177/0049124109335735>
- ZIMMERMANN, T. (2020): "Social Influence or Rational Choice? Two Models and Their Contribution to Explaining Class Differentials in Student Educational Aspirations", *European Sociological Review*, 36(1), pp. 65-81. <https://doi.org/10.1093/esr/jcz054>



